

Principy digital humanities

1. Úvod: Digitální metody, digitální nástroje a vizualizace dat

Jindřich Marek

Plán výuky (2024)

1. Úvod: Digitální metody, digitální nástroje a vizualizace dat. 22. 2.
(~~1.3.~~)
2. „Kulturní data“. 8. 3.
3. Práce s digitálními daty. 15. 3.
4. Vizualizace vědeckých dat. 22. 3.
5. Analýza textů. 5. 4.
6. Sociální data a sociální sítě. 12. 4.
7. Digitální mapy. 19. 4.
(~~26.4.~~)
8. Velké jazykové modely a jejich využití. 3. 5.
9. Umělá inteligence, umělá estetika. 10. 5.
10. Game studies a rozšířená realita, digitální materialita. 17. 5.

Moodle

- <https://dl1.cuni.cz/course/view.php?id=5734>

Témata

- vymezení: digitální obsah – nástroje – metody
- digitální metody
 - charakteristika, přístupy, methodological commons
- typický cyklus výzkumných aktivit
- digitální kulturní data (obsah) a metody
- příklady specifických metod
 - cultural analytics, distant reading, SNA, MCA
- digitální nástroje a vizualizace dat

Digital humanities

- digitální nástroje
 - staré metody
 - nové možnosti vyhodnocení
- digitální studia jsou ale ještě něco navíc
- interaktivita

Základní literatura ke kurzu

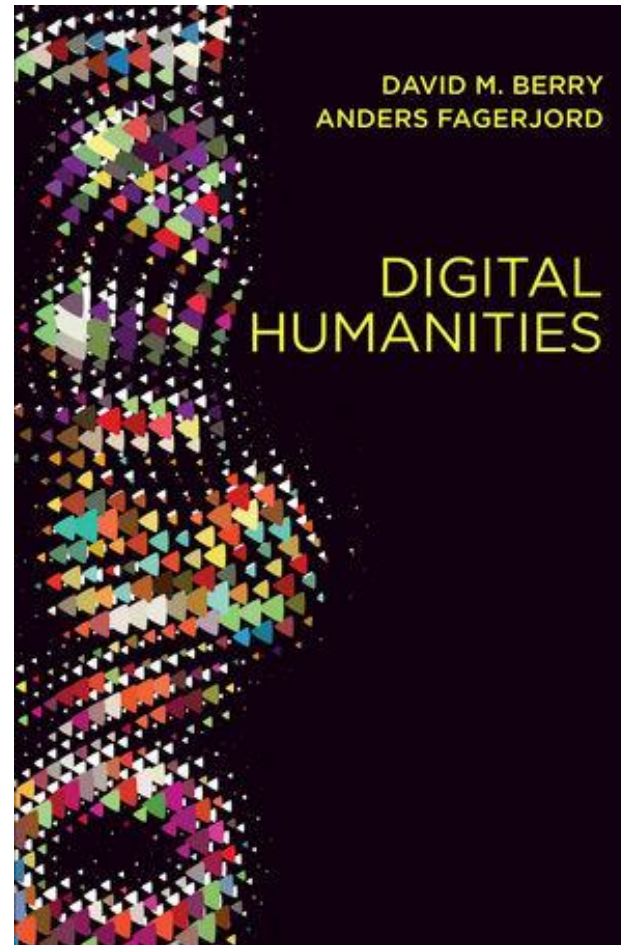


A NEW COMPANION TO

DIGITAL HUMANITIES

EDITED BY
SUSAN SCHREIBMAN, RAY SIEMENS,
AND JOHN UNSWORTH

WILEY Blackwell



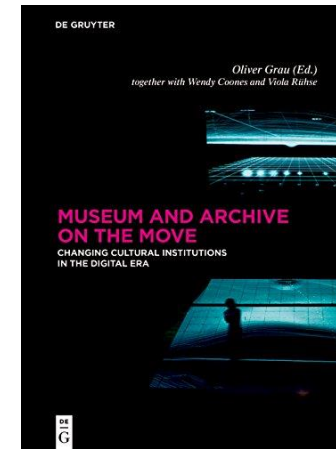
DAVID M. BERRY
ANDERS FAGERJORD

DIGITAL HUMANITIES



Between
Humanities and
the Digital

edited by
Patrik Svensson and
David Theo Goldberg



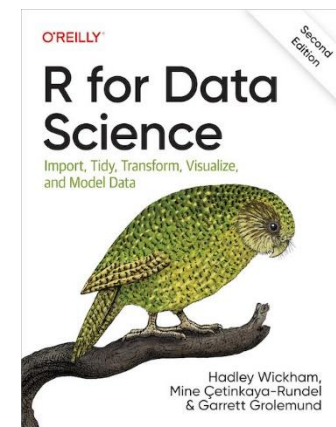
DE GRUYTER

Oliver Grau (Ed.)
together with Wendy Coones and Viola Rabus

MUSEUM AND ARCHIVE ON THE MOVE

CHANGING CULTURAL INSTITUTIONS
IN THE DIGITAL ERA

018



O'REILLY

R for Data Science

Import, Tidy, Transform, Visualize,
and Model Data

Second
Edition

Hadley Wickham,
Mine Cetinkaya-Rundel
& Garrett Grolemund

Digital humanities v českých zemích

- portál Knihovny Akademie věd
 - <https://digitalhumanities.cz>
- Česká asociace pro digitální humanitní vědy
 - <https://www.czadh.cz> (nepříliš časté příspěvky)
- Centrum pro digital humanities FF UK
 - <https://cdh.ff.cuni.cz>

Digital Humanities

Zpřístupňování, uchovávání a záchrana pramenů v digitálním věku

S podporou Strategie AV21 a programu Paměť v digitálním věku

Aktuálně

ZOBRAZIT VŠE

POUZE AKCE

POUZE AKTUALITY

POUZE NEZAŘAZENÉ

21. 11. 2023

Understanding, Meaning, and Representations in Large Language Models

Significant in contemporary language processing applications, Large Language Models (LLMs) have become indispensable across diverse domains, ranging from improving chatbot interactions and language translation to code generation and the creation of imaginative...

Termín a místo konání:
4.12. 2023 (14:00), MFF UK, Malostranské nám.
25, 4.patro, místnost S1

AKCE

23. 10. 2023

Seminar of formal linguistics

The program of Linguistic Mondays for the rest of October.

Termín a místo konání:
23/10/2023 and 30/10/2023, Faculty of Mathematics and Physics, Malostranské nám.
25, 4th floor, room S1 (no. 428)

AKCE

26. 9. 2023

Workshop: How Research Infrastructure Empowers SSH Researchers

Workshop bude pojednávat o tom, jak DARIAH přispívá k výměně znalostí a spolupráci, přičemž poskytuje platformy (SSH Open Marketplace a DARIAH Campus), které umožňují výzkumným pracovníkům najít potřebné nástroje SSH a školicí materiály.

Termín a místo konání:
2. října 2023, 14:00 - 15:30, Institute of Czech Literature of the CAS, Na Florenci 1420/3, Prague, Lower Conference Hall

AKCE

26. 9. 2023

DARIAH editathc

SSH Open popisu me DARIAH.

Nabídka D services/ti

Termín a n 9 - 16.10.

DETAIL
























DETAIL

DETAIL

Seznam členů

Členstvo

Filter Sort

 <p>Anna Chromá Institute of Czech Language and Theory of Communication amacroma@ff.cuni.cz</p> <ul style="list-style-type: none">linguisticsdevelopmental psychologyexperimental researchcorpus linguisticsCoCReLAR → eye-trackingexperimental design <p>V rámci doktorského studia (FF UK) jsem svůj disertační výzkum ovsprávním zájmem a sociální kognicí...</p> CoCReLA Corpora of Czech as the First Language in Acquisition → Lindat	 <p>Jan Chromý Institute of Czech Language and Theory of Communication evcl.ff.cuni.cz jan.chromy@ff.cuni.cz</p> <ul style="list-style-type: none">linguistics → psychologyweb-based experimentsdata analysisR → PC Bee Farm <p>Jan Chromý je docentem na Ústavu českého jazyka a teorie komunikace FF UK, jeho hlavní výzkumné zájmy:</p> Úvod do práce s jazykovými korpusy Empirické metody v lingvistice	 <p>Václav Cvrček Institute of the Czech National Corpus wiki.korpus.cz/wiki/vacv.cvrcek@ff.cuni.cz</p> <ul style="list-style-type: none">linguisticscorpus linguisticsR → Shiny → Perl Práce s korpusy: případové studie Úvod do práce s jazykovými korpusy	 <p>Lucie Doležalová Institute of Greek and Latin Studies cuni.academia.edu/Lucie_Dolezalova lucie.dolezalova@ff.cuni.cz</p> <ul style="list-style-type: none">philologymedieval studies → historyliteraturedigital editionsMedieval Latin <p>Latinská medievalistka se zájmem o rukopisnou kulturu pozdního středověku...</p> A4 → Lipnická bible Opuscula	 <p>Ondřej Fůsik Department of English Language and ELT Methodologies uagf.ff.cuni.cz/ona_fusik/ondrej.fusik@ff.cuni.cz</p> <ul style="list-style-type: none">linguistics → philologytheory → historylanguage analysisdigital ethnographydigital humanitiesMedieval LatinOld EnglishModern Englishhistorical linguisticscorpus linguisticsparallel corporaCC0 → PythonTranskript → LLM <p>Ve svém doktorském studiu se zabývám jazykovou dynamikou mezi latinou a starou angličtinou. Má hlavně:</p> Bosworth-Toller → Lindat Lipnická bible → A4 TEI Cookbook → Opuscula → ONCO Introduction to Digital Humanities Digitalní metody a AI v humanitách a sociálních vědách	 <p>Martin Hanker Institute of Asian Studies martinhanker.com/ mh@martinhanker.com</p> <ul style="list-style-type: none">linguistics → philologyanthropologylanguage analysisdigital ethnography <p>V obecné rovině se věnuji právní, literární a hudební produkci Tibetu, konkrétně pak písmo tibetské...</p> → Lindat A4 TEI Cookbook → Opuscula	 <p>Michal Křen Institute of the Czech National Corpus wiki.korpus.cz/wiki/mkren/ michal.kren@ff.cuni.cz</p> <ul style="list-style-type: none">linguistics → informaticsquantitative linguisticscorpus linguisticsvariability in languagecollocationslanguage changePerl → LaTeX → SQLCC0 → Linux <p>Jsem absolventem informatiky na MFF UK (Mgr.) a matematické lingvistiky na FF UK (Ph.D.). Jsem hlavně:</p> Práce s korpusy: případové studie Úvod do práce s jazykovými korpusy	 <p>Karolina Lochmanová Institute of Translation Studies karol.stma@gmail.com</p> <ul style="list-style-type: none">translation → linguisticsliteratureDoktorandská Ústavu translatologie, výzkum prelaadů české literatury do španělštiny a tvorba bibliografie...
 <p>Eva Maria Luef Department of English Language and ELT Methodology uagf.ff.cuni.cz/~luef/ EvaMaria.Luef@ff.cuni.cz</p> <ul style="list-style-type: none">psycholinguisticssociolinguisticssecond language acquisitionlanguage evolutionweb-based experimentsdata analysisnetwork scienceR → Gephi → OpenAI API <p>Eva Maria Luef je docentkou anglické lingvistiky. Její hlavní výzkumné oblasti jsou psycholingvistika...</p> Language Network Analysis	 <p>Jindřich Marek Institute of Information Studies and Librarianship uagf.ff.cuni.cz/~jmarek/ jindrich.marek@ff.cuni.cz</p> <ul style="list-style-type: none">book historysocial historymedieval studiesdata explorationnetwork science → LLMPython → R → LaTeXGephi → OpenAI APIXML <p>Dlouhodobě se zabývám rukopisnou kněžní kulturou v českých zemích ve střešování a v našem novověku...</p> Kněhověda.cz	 <p>Anna Michalová Institute of Czech Language and Theory of Communication Czech Academy of Sciences, Czech Language Institute ujc.cvcz.cz/ak...html a.michalova@ujc.cas.cz</p> <ul style="list-style-type: none">linguistics → philologydigital editionsautomatic text recognitionnetwork science → LLMPython → R → LaTeXGephi → OpenAI APIXML <p>Old Czech → Middle Czech</p> Transkript → Pars OCR Jsem doktorkou na Ústavu českého jazyka a teorie komunikace Filozofické fakulty University Karlovy... A4 → najdišlov.cz	 <p>Jiří Milíčka Institute of the Czech National Corpus wiki.korpus.cz/wiki/jiri_milicka/jiri.milicka@ff.cuni.cz</p> <ul style="list-style-type: none">linguisticsquantitative linguisticscorpus linguistics → LLMstylistometry → ArabicCzechOpenAI API → Python → RDelphi → Pascal → LaTeXNeat → GraphvizGephi <p>Kvantitativní a korpusový lingvistika, momentálně nadány z velkých jazykových modelů. Svého času jsem:</p> Práce s korpusy: případové studie Úvod do práce s jazykovými korpusy Využití velkých jazykových modelů	 <p>Světlana Šrnidová Müllerová Department of English Language and ELT Methodology facebook.com/srnidovamullerova/ https://uagf.ff.cuni.cz/o-ustave/externisti/svetlana-smidova-mullerova/ https://cuni.academia.edu/SvetlanaSmidovaMullerova svetlana.mullerova@ff.cuni.cz</p> <ul style="list-style-type: none">linguistics → philologymedieval studiescontact linguisticslexical borrowingOld Englishhistorical linguisticscorpus linguisticsdigital editionsCC0 → TEI → HTML <p>Ve svém doktorském studiu se zabývám anglo-standardskému jazykovému kontaktu, zejména se zaměřením n...</p> Bosworth-Toller Lipnická bible → ONCO	 <p>Karel Pacovský Institute of Greek and Latin Studies cuni.academia.edu/Karel_Pacovsky/ karel.pacovsky@ff.cuni.cz</p> <ul style="list-style-type: none">medieval studiestheory → philologyMedieval Latin → literaturehistorydigital editions → Lindat A4 Lipnická bible → Opuscula	 <p>Zaneta Pivozová Institute of Czech Language and Theory of Communication facebook.com/zaneta.pivozova zan.pivoz@gmail.cz</p> <ul style="list-style-type: none">linguistics → lawphilologyDoktorandská Ústavu českého jazyka a teorie komunikace FF UK v oboru Český jazyk, v disertační práci ...	 <p>Martin Roček Institute of Greek and Latin Studies facebook.com/martin.rocek/ martin.rocek@ff.cuni.cz</p> <ul style="list-style-type: none">philology → historydigital humanitiesdictionaries → LLMmachine learningdigital editionsPython → programmingJS → CSS → HTMLNeSQL → SQL → LinuxHTBAdobe Creative Cloud → 4 <p>Ve svém doktorském studiu se věnuji středověkým mýtům (Gjerdnesen Semmel Davidsen) a snům obecně...</p> Bosworth-Toller Image Annotation Tool A4 → Scribium → Lindat → najdišlov.cz Lipnická bible TEI Cookbook → Opuscula Introduction to Digital Humanities
 <p>Tomáš Šavčenko Department of English Language and ELT Methodology Department of Linguistics tsavcenko@gmail.com</p> <ul style="list-style-type: none">linguisticsnetwork analysis → LLMbig data → corpussemanticsPython → JavaScriptPHP → GitHubJupyter Notebooks → R <p>Jsem magisterským studentem anglického jazyka a obecné lingvistiky. Obecně mě zajímá propojení počítač...</p>	 <p>Michal Šarabál Institute of the Czech National Corpus wiki.korpus.cz/wiki/msharab/ michal.sarabal@ff.cuni.cz</p> <ul style="list-style-type: none">corpus linguisticscontrastive linguisticslexicographydictionariescrowdsourcinglexicography → corpusPascal <p>Mgr. Michal Šarabál, Ph.D. (1979), absolvent FF UK, Mgr., Český jazyk a literatura – loyblina 05</p> Práce s korpusy: případové studie Úvod do práce s jazykovými korpusy	 <p>Pavel Štichauer Department of Romance Studies urs.ff.cuni.cz/~pshauer/ pavel.stichauer@ff.cuni.cz</p> <ul style="list-style-type: none">linguistics → morphologyword formationcorpus linguisticsdata analysis → big dataR → LaTeX <p>Morfologie románských jazyků, zvláště italský a jejich dialekty. Slovozná teorie, teorie flex...</p> The MORPAR project: A quantitative approach to mixed perfective auxiliary systems in Italo-Romance	 <p>Ondřej Tichý Department of English Language and ELT Methodologies uagf.ff.cuni.cz/tichy/ ondrej.tichy@ff.cuni.cz</p> <ul style="list-style-type: none">linguistics → philologyhistorical linguisticslexicographydigital humanitiesprogrammingdata analysis → KonTextCC0 → SQL <p>Historická a korpusová lingvistika se zaměřením na anglický jazyk, kvantitativní a komputační lingvi...</p> Bosworth-Toller najdišlov.cz → MeCz Elektronické mluvnice současné angličtiny Data Processing Introduction to Digital Humanities Digitalní metody a AI v humanitách a sociálních vědách English Diachronic Corpora	 <p>Martina Vacková Reiterová Institute of World History cuni.academia.edu/Martina_Vackova_Reiterova/ martina.reiterova@ff.cuni.cz</p> <ul style="list-style-type: none">history → sociologyrhetoricsdiscourse analysisdata analysisSQL → Python → ATLAS.tI <p>Ve svém výzkumu se zabývám analýzou reprezentativních strategií obranných hnutí v tzv. keltických z...</p>	 <p>Jitka Veroniková Institute of Phonetics ff.cuni.cz jitka.veronikova@ff.cuni.cz</p> <ul style="list-style-type: none">phoneticsspeech communicationCzech for foreignerssecond language acquisitionCzech <p>Působím ve Fonetickém ústavu FF UK. Věnuji se zvukové stránce jazyka a řeči se zaměřením na češtinu...</p> ProCafee (Pronunciation of Czech for Foreigners) – Vyslovnost češtiny pro cizince	 <p>Adrian Jan Zalina Institute of Czech and Deaf Studies researchgate.net/profile/AdrianJanZalina/ adrian.zalina@ff.cuni.cz</p> <ul style="list-style-type: none">linguisticsdata-driven languagecorpus linguisticsparallel corporavariability in languagecorpus-assisted discourse studiesCzech → Polishgrammarsecond language acquisitioncontrastive linguistics → 1TEI/OK → Sketch EngineKonText <p>Je absolventem magisterského studia slavistiky a doktorského studia korpusové lingvistiky Akademick...</p> Korpusy slovanských jazyků Korpusová analýza jazykových textů	

Digitální metody

DH: prudký rozvoj v posledních 20 letech

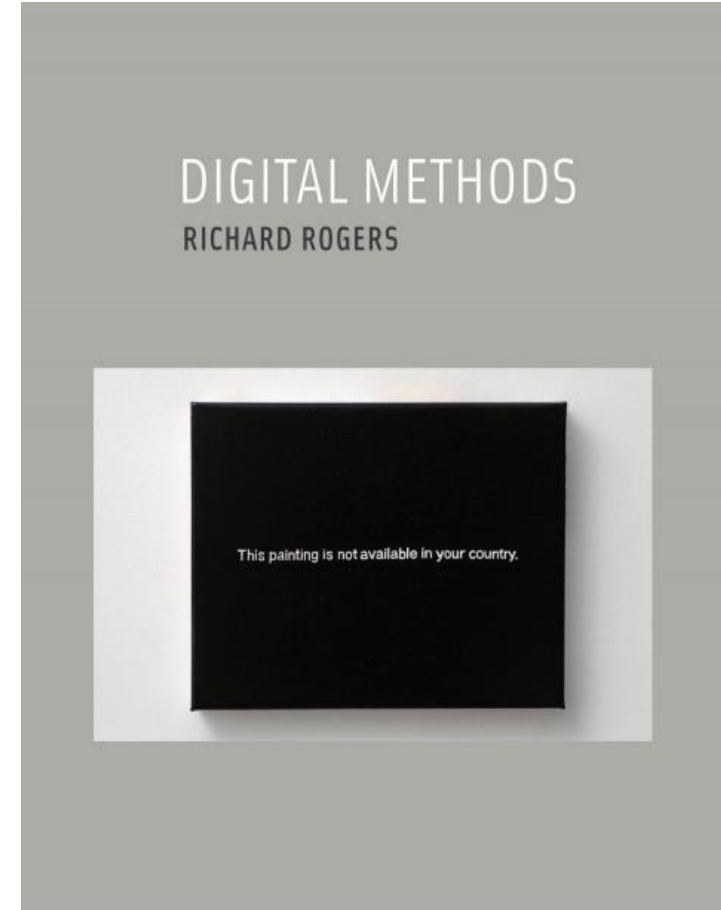
- mnohem více digitálních sbírek online
 - primární pramenný materiál pro výzkum dějin literatury, historie, jazykovědy, klasických studií, hudební a filmové vědy a dalších disciplín
- rozvoj webových digitálních nástrojů a přístupů
 - vytváření, analýza a sdílení digitálního výzkumu
 - interaktivita, kultura sdílení (participatory culture)
- důraz na spolupráci humanitních a technických (případně přírodních) věd

Digital humanities

- digitální obsah
- digitální nástroje
- digitální metody

Digital humanities

- digital humanities
- humanities+ (Liu)
- digital methods (Rogers)



Digital humanities: teorie

- teorie spojená s vytvářením počítačových systémů je implicitní, navíc nastala doba „post-teoretická“
- je třeba zvládnout záplavu digitálních dat, která vznikla digitalizací nebo již jen v digitální podobě
- někteří badatelé se domnívají, že teorie není třeba a postačují data
- epistemologická a metodická inovace z technických a přírodovědných disciplín

McCarty: Methodological Commons (I)

- technické metody z oblastí mimo oblast humanitních věd, např. technické vědy a informatika (např. pro využití, vizualizaci a modelování digitálního obsahu)
- nové způsoby spolupráce mezi disciplínami a komunitami (partnerství s přírodovědnými, technickými obory a výzkumem kulturního dědictví)

McCarty: Methodological Commons (II)

- často kombinace datových typů, technických metod a několika technologií, například textových, databázových, obrazových dat (také videa nebo zvuku) a geografických informačních systémů (GIS)
- formální metody analýzy a návrhu zdrojových dat a modelování možných technických přístupů
- metody práce s rozsáhlými datovými zdroji, agregace materiálů z více sbírek nebo zdrojů

DARIAH-DE

Research activities

1 *Capture*

conversion
data recognition
discovering
gathering
imaging
recording
transcription

2 *Creation*

designing
programming
web development
writing

3 *Enrichment*

annotating
cleanup
editing

4 *Analysis*

content analysis
network analysis
relational analysis
spatial analysis
structural analysis
stylistic analysis
visualization

5 *Interpretation*

contextualizing
modeling
theorizing

6 *Storage*

archiving
identifying
organizing
preservation

7 *Dissemination*

collaboration
commenting
communicating
crowdsourcing
publishing
sharing

Digitální data


- „kulturní data“


Kultura

- sociální chování a normy v lidských společnostech
- materiální kultura = fyzická vyjádření kultury
 - technologie, architektura, umění

World Digital Library

- <https://www.loc.gov/collections/world-digital-library/about-this-collection/>

Original Format	
Photo, Print, Drawing	13,300
Book/Printed Material	3,952
Map	1,422
Manuscript/Mixed Material	175
Periodical	45
Film, Video	11
Newspaper	11
Notated Music	11
Audio Recording	9
Web Page	1

Online Format	
Image	18,896
Compressed Data	5,069
PDF	3,990
Online Text	541
Video	11
Audio	9
Web Page	2

COLLECTION

World Digital Library

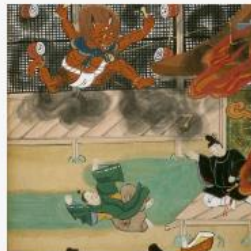
[About this Collection](#)

[Collection Items](#)

Featured Content



Morning Glory Flowers:
National Diet Library



The Origin of Tenjin:
National Diet Library



Annals of Creation: National
Library of China



One Hundred Aspects of
the Moon: Moon and
Smoke: National Diet ...



Itinerary Book Kept During
the journey to East India:
National ...



About this Collection

[Rights and Access](#)

Expert Resources

[Main Reading Room](#)

[General Collections](#)

[Ask a Librarian](#)

[African & Middle Eastern Reading Room](#)

[Ask a Librarian](#)

[Listen to this page](#)

About this Collection

This collection contains cultural heritage materials gathered during the World Digital Library (WDL) project, including [thousands of items contributed by partner organizations](#) worldwide as well as content from Library of Congress collections. The original World Digital Library site (preserved in [LC's Web Archives here](#)) and all descriptive metadata were translated from English and made available in six additional languages: [Spanish](#), [Portuguese](#), [French](#), [Arabic](#), [Russian](#), and [Chinese](#). All item records include narrative descriptions submitted by the contributing partners and enhanced by WDL researchers to contextualize the item and its cultural and historical importance. Books, manuscripts, maps, and other primary materials in the WDL collection are presented in their original languages; more than 100 languages are represented, including many lesser known and endangered languages.

Europeana (Collections)

- <https://www.europeana.eu/cs>
- kulturní dědictví (nejen dokumentární)

Poskytujeme vám přístup k milionům předmětů kulturního dědictví z institucí napříč Evropou. Objevte umělecká díla, knihy, hudbu a videa o umění, novinách, archeologii, módě, vědě, sportu a mnoho dalšího.



[29 734 524](#)

OBRÁZEK



[21 984 501](#)

TEXT



[771 193](#)

ZVUK



[527 749](#)

VIDEO



[17 723](#)

3D



Email courses

Subscribe to our email courses to learn something new in 2022

EXPLORE

Nové příběhy



VÝSTAVY

Women writing birds

Learn about pioneering American ornithologists and their early environmental activism

Kaunas Esch-sur-Alzette Novi Sad

PŘÍSPĚVKY NA BLOGU

European Capital of Culture

Discover the 2022 European Capital of Culture cities: Esch-sur-Alzette, Kaunas and Novi Sad



PŘÍSPĚVKY NA BLOGU

Wilhelm Dreesen

Discover a pioneering 19th century German photographer who was a globetrotter and influencer



PŘÍSPĚVKY NA BLOGU

Lilly Steinschneider

Hungary's first female pilot who conquered the skies in 1912



PŘÍSPĚVKY NA BLOGU

Abdullah Frères

Photographer brothers who documented life and people in the late Ottoman Empire



VÝSTAVY

Louise Weiss

Angažovaná Evropanka



PŘÍSPĚVKY NA BLOGU

Ismail al-Jazari

The Muslim inventor who may have inspired Leonardo da Vinci



PŘÍSPĚVKY NA BLOGU

Augusta and Anna Curiel

Photographer sisters who chronicled early 20th century life in Suriname

Zpětná Vazba

Google Arts & Culture

<https://artsandculture.google.com>



M
O

Musée d'Orsay, Paris
Paris, Francie



MoMA

MoMA The Museum of
Modern Art
New York, Spojené státy



V
M

Van Gogh Museum
Amsterdam, Nizozemsko



National
Gallery
of Art

National Gallery of Art,
Washington DC
Washington, DC, Spojené státy



LE GALLERIE
D'ART UFFIZI

Uffizi Gallery
Firence, Itálie



A
I

The Art Institute of Chicago
Chicago, Spojené státy



NATIONAL
MUSEUM OF
NATURAL
HISTORY

Smithsonian National
Museum of Natural History
Washington, DC, Spojené státy



THE
MET

The Metropolitan Museum of
Art
New York City, Spojené státy



MUSEO SOROLLA

Sorolla Museum
Madrid, Španělsko



RIJKSMUSEUM

Rijksmuseum
Amsterdam, Nizozemsko



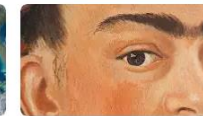
E
T

Eiffel Tower
Paris, Francie



NASA

NASA
Washington, DC, Spojené státy



EL OLMEDO

Museo Dolores Olmedo
Mexico, Mexiko



The State Hermitage
Museum

The State Hermitage
Museum
St. Peterburg, Rusko



Mh Mauritshuis

Mauritshuis
Den Haag, Nizozemsko

Meet an ancient animal

Get up close with Virtual Reality



Amurosaurus

Play on the app
 Take photos with Art Transfer, find your Art Selfie, and more

Museum explorer

Where do you want to visit today?

Přeložit pomocí Google

PROZKOUMAT	PROZKOUMAT	PROZKOUMAT	PROZKOUMAT
BALTIC Centre for Contemporary Art Vienna	Museum of Fine Arts Vienna	Dr. Bhau Daji Lad Museum Mumbai	Australian Museum New South Wales

Today's top picks

Zoom Into the Ripley Discover the details with the B...	How Long is The Great And 9 other things to know about	3 Buildings With Amazing Stories Discover the history behind these iconic sites	One Out ... can we actually eat?	These Famous Works? Art History

Využití a (re)prezentace digitálních dat

- prostor a čas
- vizualizace informací
- propojená data (linked data) a ontologické metody
- vytváření a vývoj digitálních sbírek pro výzkum
- využívání velkých sbírek textů
- vytváření digitálních edic

Vybrané metody DH

- cultural analytics (Manovich)
- social network analysis
- distant reading (Moretti)
- correspondence analysis
- ...

Cultural analytics (Manovich)

- vytvoření datasetu
- automatická analýza založená na vizuálních prvcích
- vizuální uspořádání generovaných výsledků
 - analýza selfíček na Instagramu: <http://www.selfiecity.net/>
- proti zakládání teorií na malých (a/nebo výběrových) datasetech (např. Italská renesance apod.)

Selfiecity investigates the Instagram self-representations in six cities around the world

- News
- About
- Projects
- Publications
- People
- Mailing List
- Twitter
- Flickr
- RSS

Uložit Tweet

The SELFIEEXPLORATORY is part of SELFIECITY

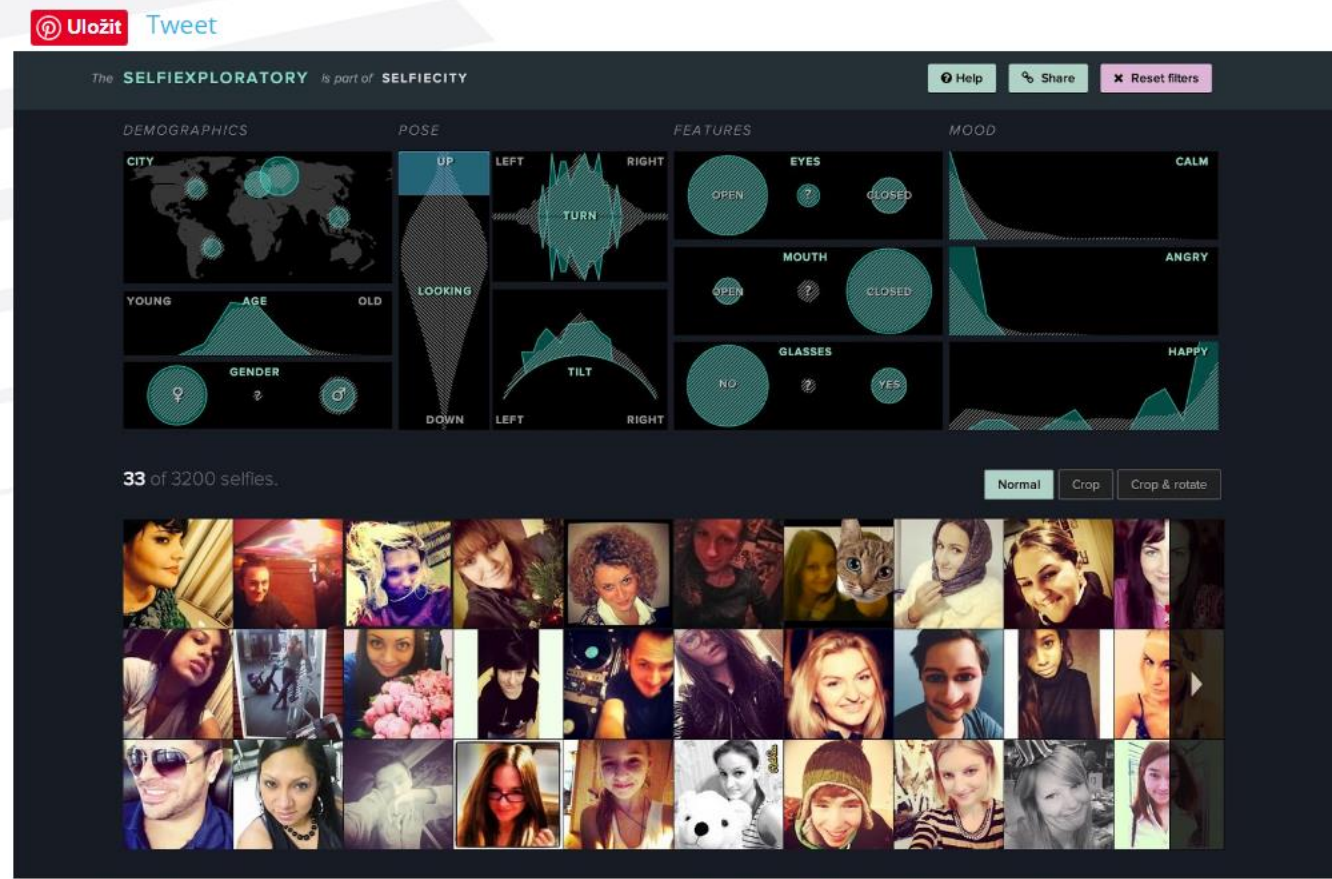
Help Share Reset filters

DEMOGRAPHICS POSE FEATURES MOOD

CITY AGE GENDER LOOKING TURN TILT EYES MOUTH GLASSES CALM ANGRY HAPPY

33 of 3200 selfies.

Normal Crop Crop & rotate

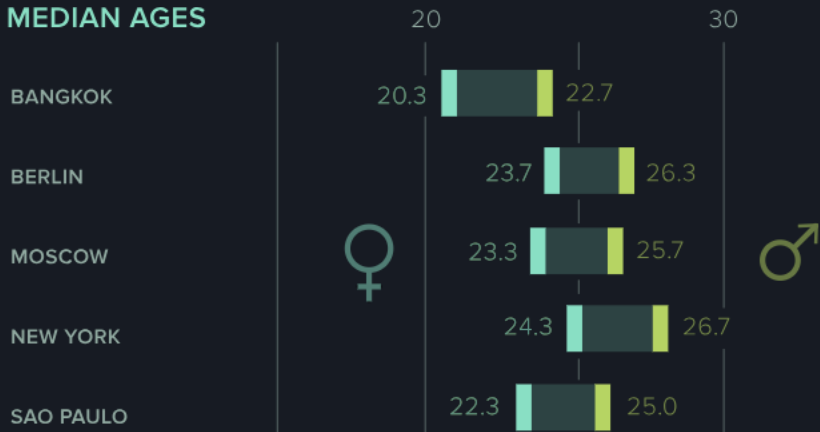
The screenshot shows the SelfieExploratory app interface. At the top, there are social media sharing options for 'Uložit' and 'Tweet'. Below that, the app title 'The SELFIEEXPLORATORY is part of SELFIECITY' is displayed along with 'Help', 'Share', and 'Reset filters' buttons. The main interface is divided into four columns: 'DEMOGRAPHICS', 'POSE', 'FEATURES', and 'MOOD'. Each column contains several interactive filters represented by circular sliders or buttons. 'DEMOGRAPHICS' includes filters for 'CITY' (a world map), 'AGE' (ranging from 'YOUNG' to 'OLD'), and 'GENDER' (with female and male icons). 'POSE' includes 'LOOKING' (UP, DOWN, LEFT, RIGHT), 'TURN' (LEFT, RIGHT), and 'TILT' (LEFT, RIGHT). 'FEATURES' includes 'EYES' (OPEN, CLOSED), 'MOUTH' (OPEN, CLOSED), and 'GLASSES' (NO, YES). 'MOOD' includes 'CALM', 'ANGRY', and 'HAPPY'. Below the filters, it shows '33 of 3200 selfies.' and a grid of 33 small selfie photos. At the bottom right of the grid, there are buttons for 'Normal', 'Crop', and 'Crop & rotate'.

A screenshot of Selfieexploratory, an interactive app we built for exploring the dataset of 3840 selfie photos. [Access the app online](#) to explore the Selfiecity dataset.

A young people's sport? Indeed.

Most people in our photos are **pretty young** (23.7 estimated median age). **Bangkok** is the **youngest** city (21.0), whereas NYC is the oldest (25.3). **Men's average age is higher** than that of women in every city. Surprisingly, more older men (30-) post selfies on Instagram than women.

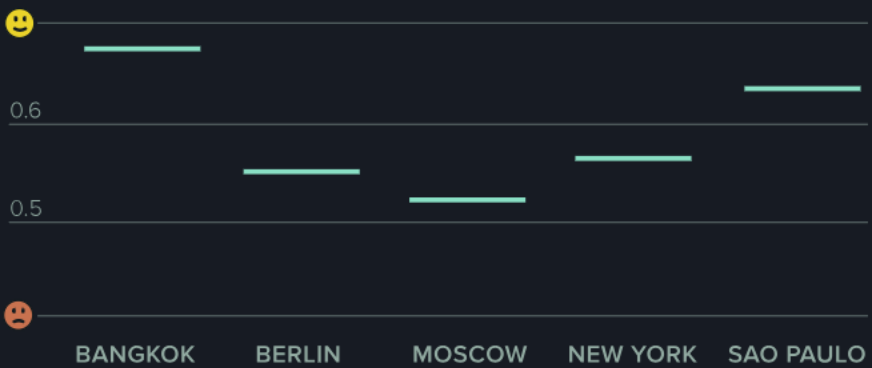
MEDIAN AGES



Bangkok, Sao Paulo are all smiles

Our mood analysis revealed that you can **find lots of smiling faces in Bangkok** (0.68 average smile score) and **Sao Paulo** (0.64). People taking selfies in **Moscow smile the least** (only 0.53 on the smile score scale).

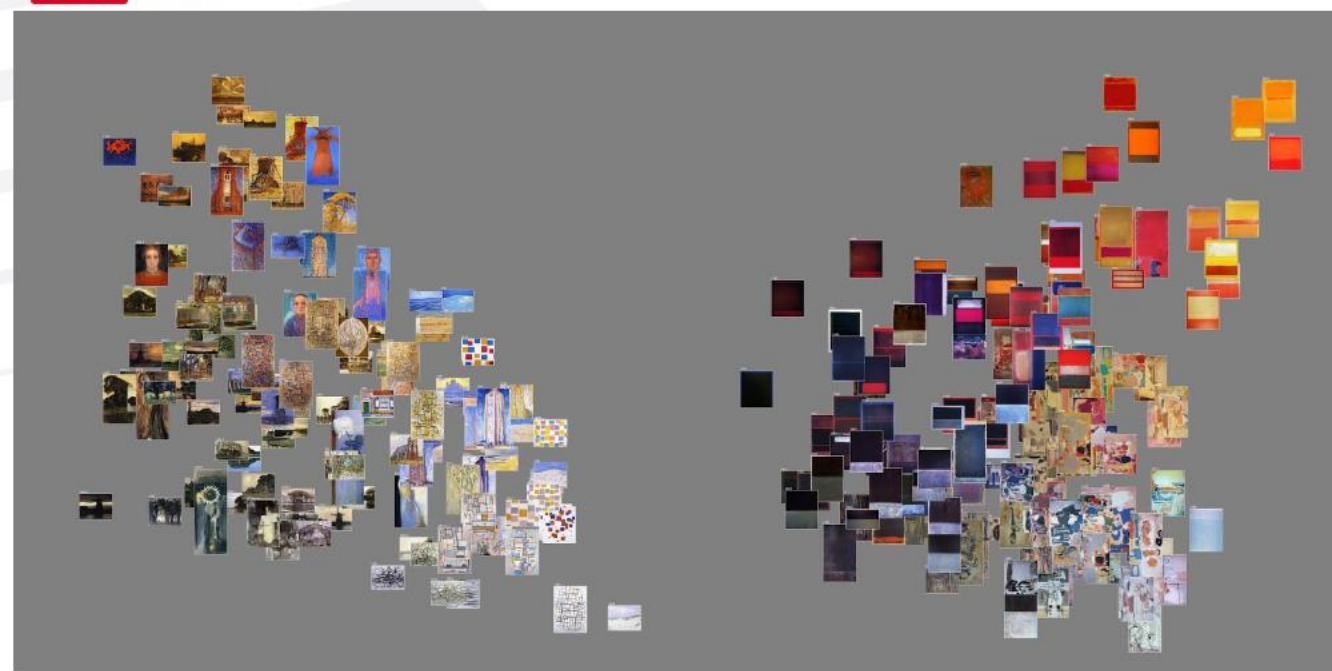
AVERAGE SMILE SCORES



Mondrian vs Rothko: Revealing the Comparative "Footprints" of the Modern Painters

- News
- About
- Projects
- Publications
- People
- Mailing List
- Twitter
- Flickr
- RSS

Uložit Tweet



Data: 128 paintings by Piet Mondrian (1905 - 1917) and 151 paintings by Mark Rothko (1944 - 1957). Mapping: X-axis: brightness mean, Y-axis: saturation mean. This visualization demonstrates how image plots can be used to compare multiple data sets. In this case, the goal is to compare similar number of paintings by Piet Mondrian and Mark Rothko (produced over comparable time periods of 13 years) along particular visual dimensions. [See the full size image on Flickr.](#)

flickr

Sign Up

Explore

Prints

Get Pro

Photos, people, or groups

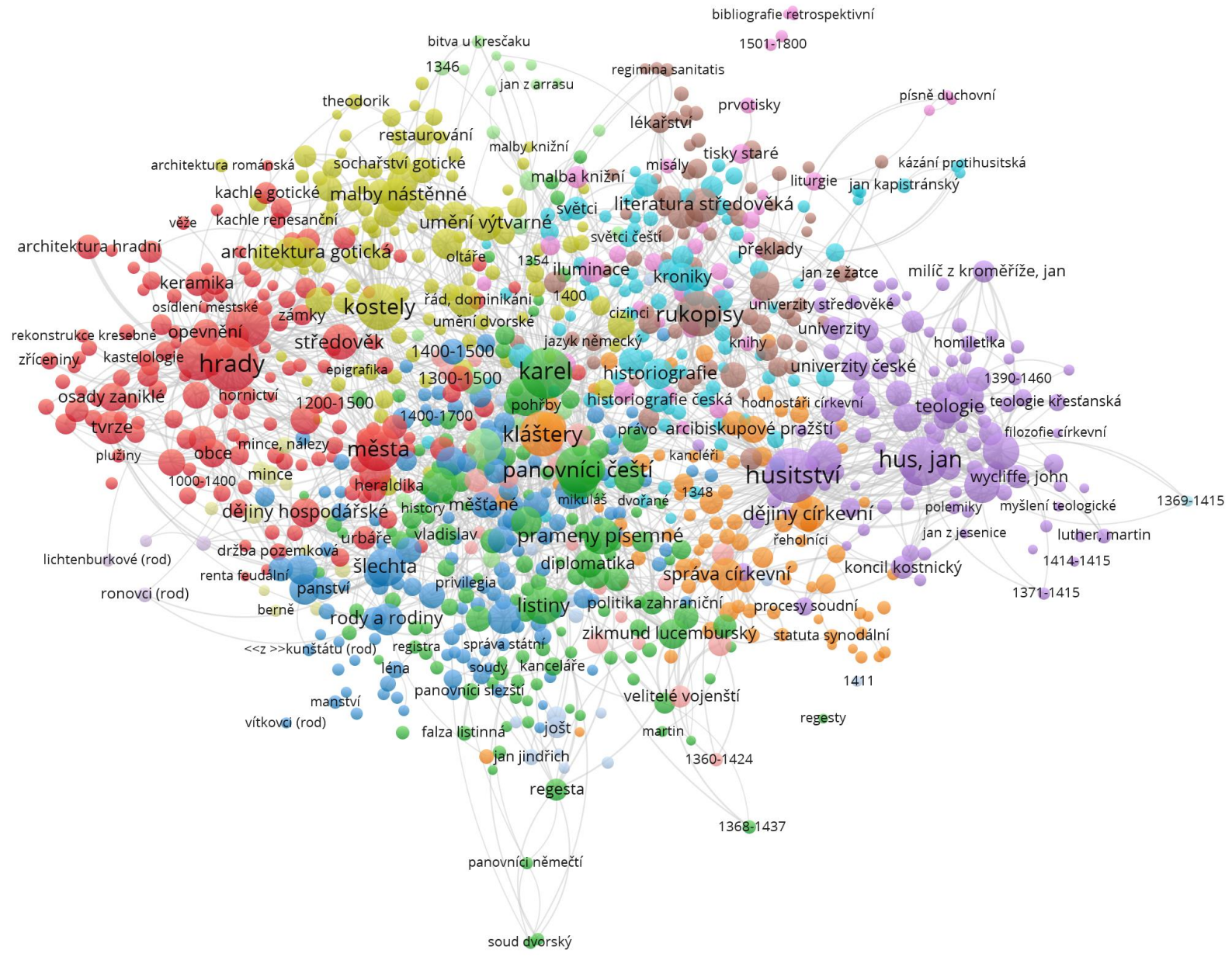


Sign In



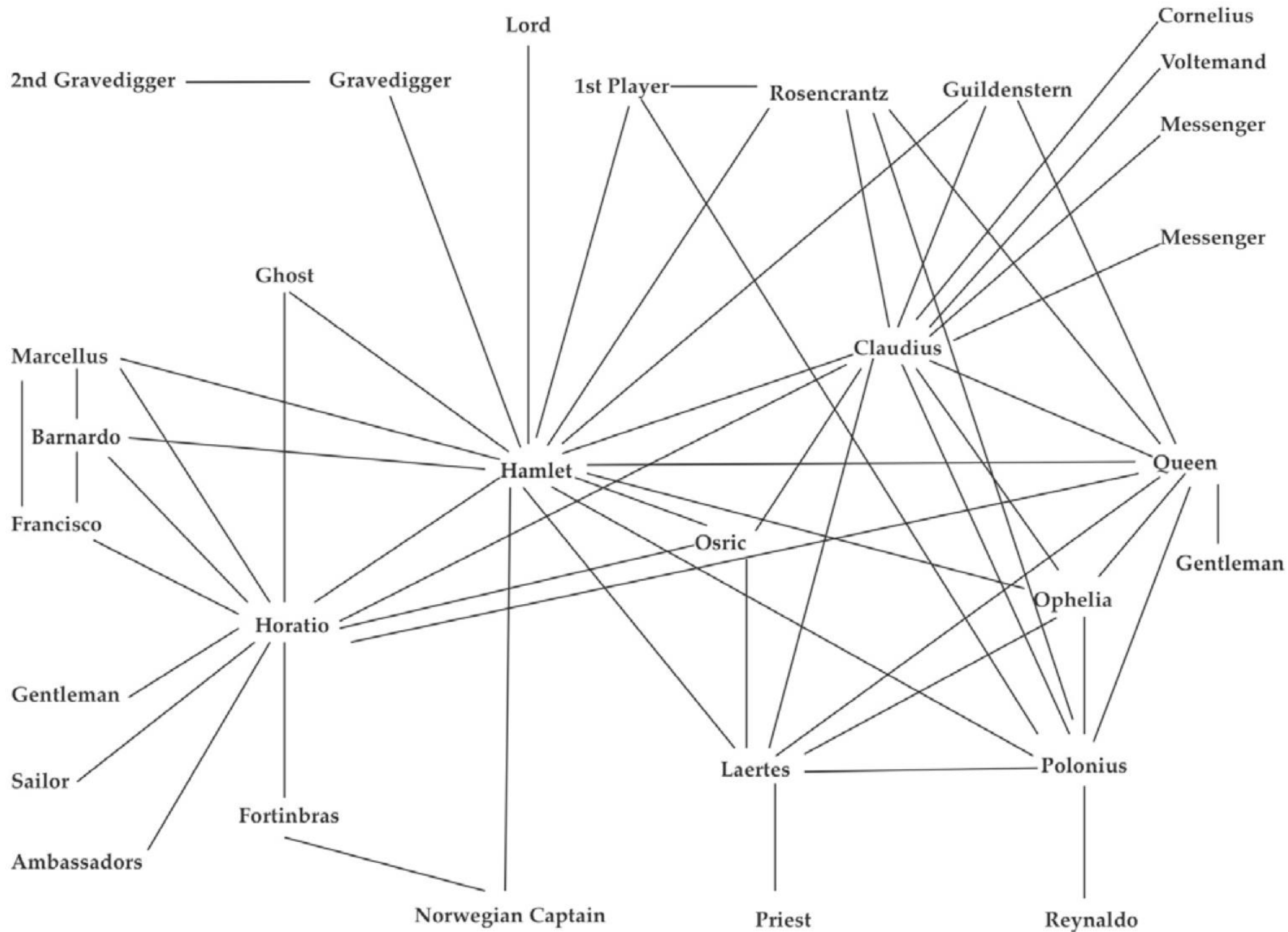
Social network analysis

- analýza sociálních sítí
- sociogram = grafické vyjádření vztahů osoby
- základní prvky: vrchol (node/vertice), hrana (edge)
- orientované, neorientované grafy (directed, undirected)
- vlastnosti grafu a jeho prvků: např. stupeň (degree), modularita (modularity), eigenvector-centrality apod.



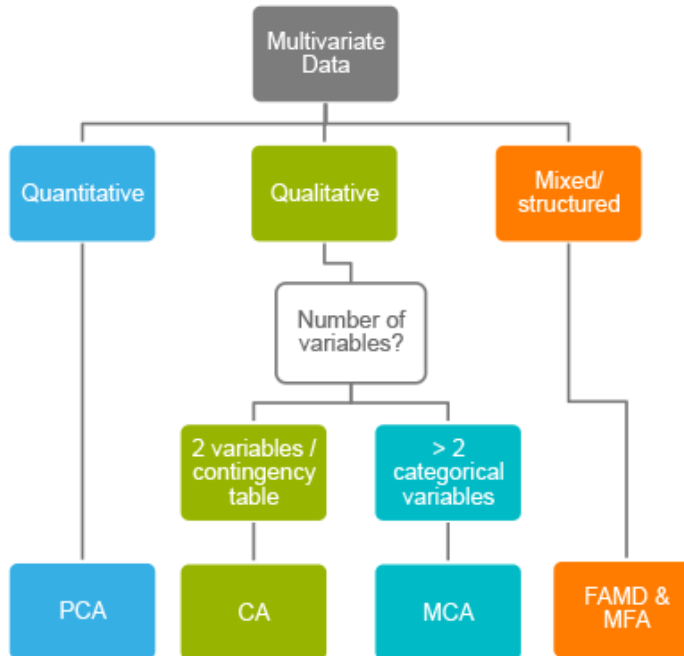
Distant reading (Moretti)

- analýza literatury ne na základě četby určitých textů, ale agregováním velkých objemů dat - počítače čtou texty nebo obrazy bez lidské interakce, lineárně, bez preferencí a strukturace
- rozpoznávání (analýza) řetězců ve velkých kulturních datech (např. sytost apod. ve snímcích obrazů)
- Metadata: zobrazení na mapě, rovnání, vizualizace
- data



Korespondenční analýza

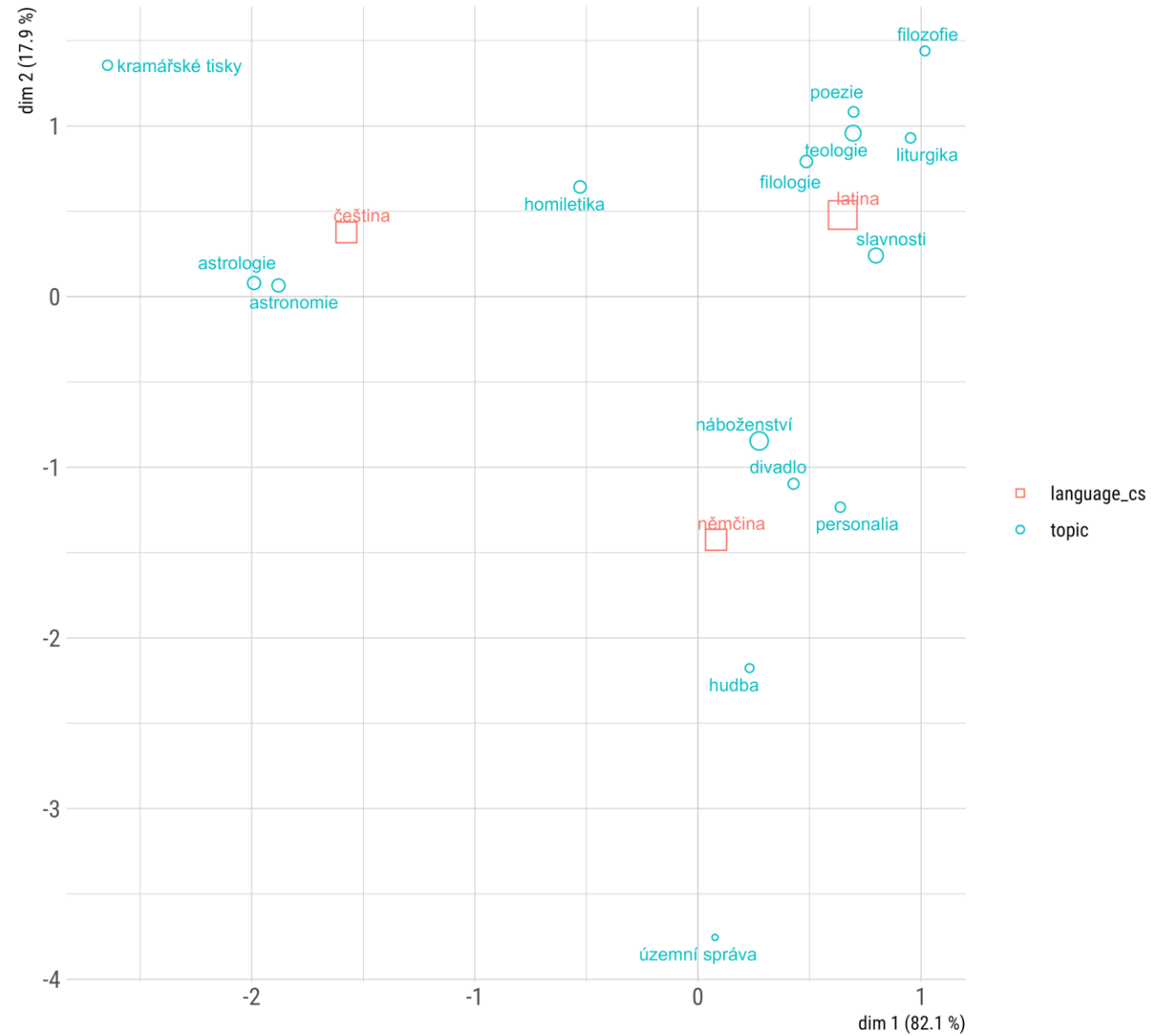
Principal Component Methods *Methods to Summarize & Visualize Multivariate Data*



- PCA: Principal Component Analysis
- (M) CA: (Multiple) Correspondence Analysis
- FAMD: Factor Analysis of Mixed Data
- MFA: Multiple Factor Analysis

Barokní knihtisk v Praze: témata a jazyky

Korespondenční analýza



Literatura

- *Cultural Analytics Lab*. Dostupné z: <http://lab.culturalanalytics.info>
- MANOVICH, Lev. *Cultural Data: Possibilities and Limitations of Digitized Archives*. In: Oliver Grau, ed., with Wendy Coones and Viola Rühse, *Museum and Archive on the Move. Changing Cultural Institutions in the Digital Era*. Berlin, Boston: De Gruyter, 2017, s. 259-276.
- MORETTI, Franco. *Distant reading*. London: Verso, 2013.
- ROGERS, Richard. *Digital methods*. Cambridge: The MIT Press, 2013.

Digitální nástroje a vizualizace dat

Vizualizace dat

- poměrně obsáhlá problematika, která zasahuje nejen do akademické, ale i např. do obchodní sféry
 - zahrnuje vytváření a studium vizuální reprezentace dat
 - cílem je jasné a efektivní sdělování informací v podobě různých druhů grafů a infografiky
 - potřeba rozumného vizuálního návrhu
- široká nabídka úzce specializovaných i komplexních nástrojů pro analýzu a vizualizaci dat, různé licence (svobodné i proprietární)

Vizuální elementy

- fyzické objekty (jako příklady)
- modely (např. skeletu)
- grafy
- tabulky
- fotografie
- kresby a diagramy

Vizualizace dat: typy vztahů (Stephen Few)

- vývoj v průběhu času
- rozdělení do tříd
- vztah části a celku
- odchylka
- rozvrstvení četnosti výskytu
- korelace
- nominální porovnání
- geografické zastoupení

Grafické vyjádření vědeckých dat

- potřeba přípravy dat pro vizualizaci
- „vhodný graf pro vhodná data“
- problematika barevných palet
- nástroje pro tvorbu grafů

Vizualizace vědeckých dat

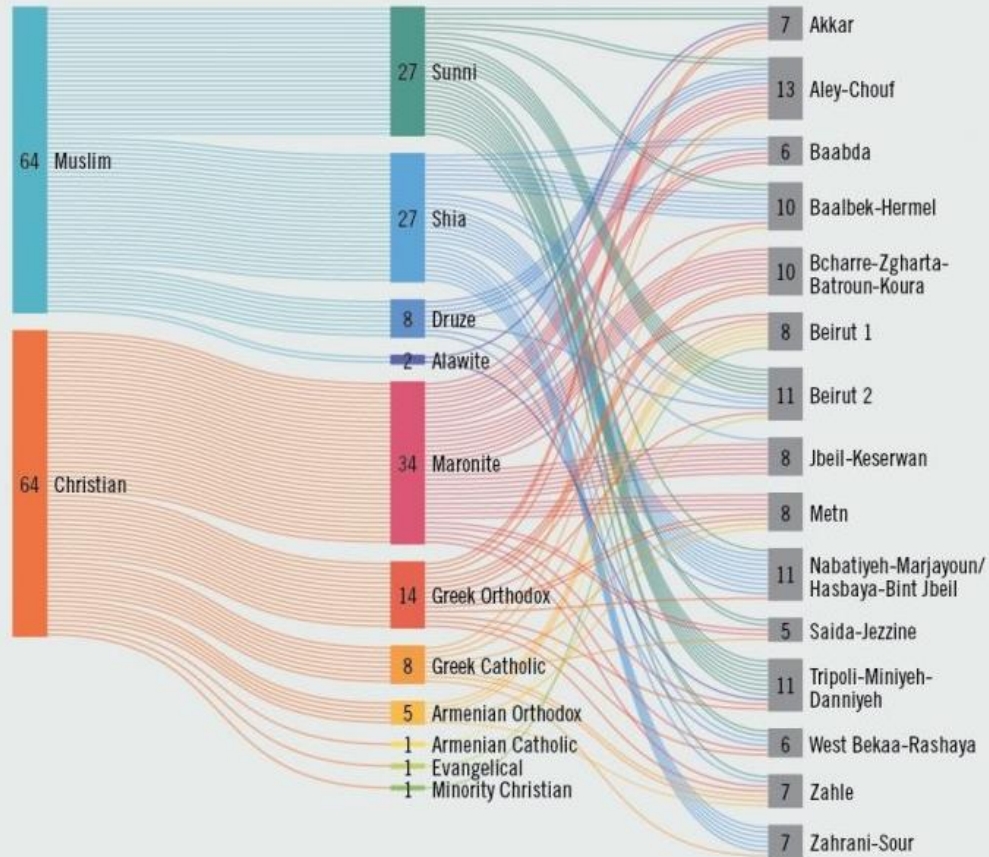
- grafické vyjádření vědeckých dat
- vyjádření dějů v prostoru
- kvantitativní analýza textů
- analýza sociálních sítí

Příprava dat pro vizualizaci

- standardizace dat
 - <https://github.com/OpenRefine/OpenRefine/wiki/Screencasts>
- nástroj pro tvorbu grafů
 - <http://rawgraphs.io/>, také běžné tabulkové procesory
- problematika barevných palet
 - <http://tools.medialab.sciences-po.fr/iwanthue/>

2017 LEBANESE ELECTION LAW

2018 SEATS BY CONFESSION
Distribution of seats by confession and district



Source: 2017 Election Law- Ministry of Interior and Municipalities © 2017 Executive.com.lb

2017 Lebanese Election Law

SUBMITTED BY
Ahmad Barclay

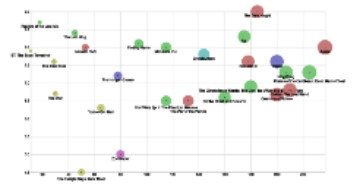
DESCRIPTION
Part of a series of visuals explaining the new Lebanese election law. The country is divided into 15 districts with semi-proportional representation, but the voting system is complicated by quotas of MPs by sect.

CHART USED
Alluvial Diagram

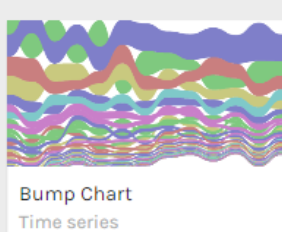
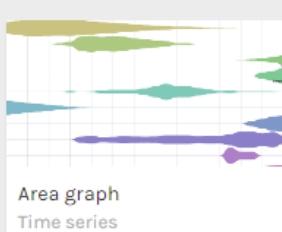
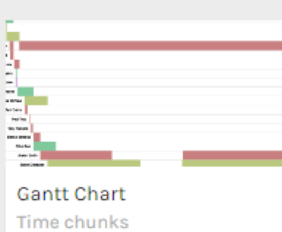
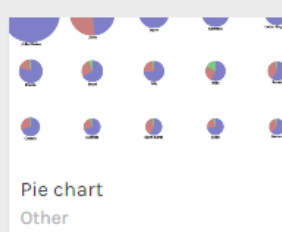
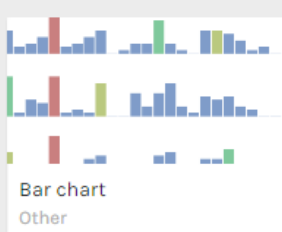
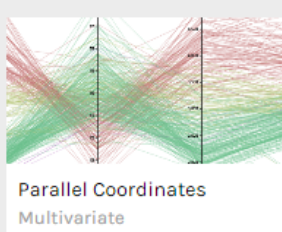
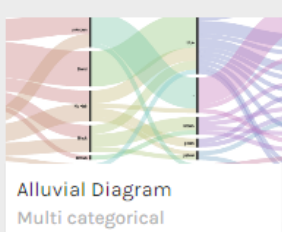
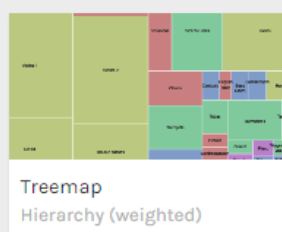
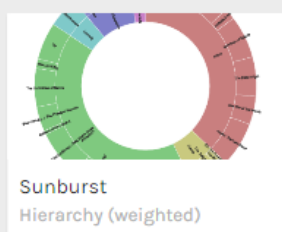
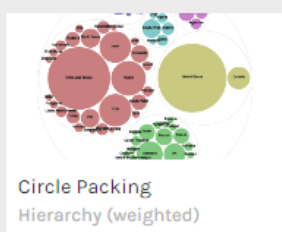
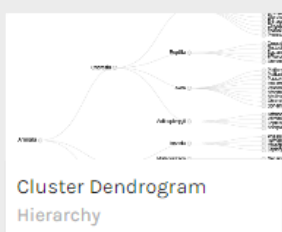
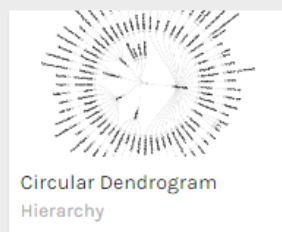
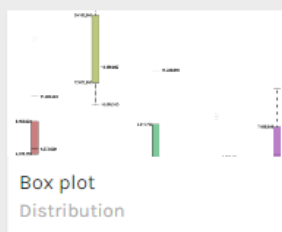
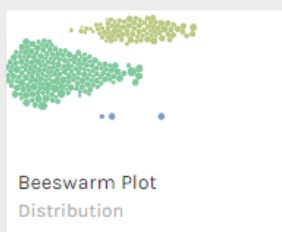
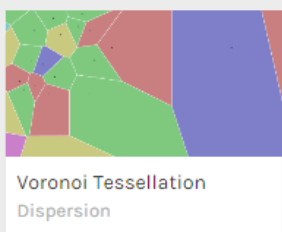
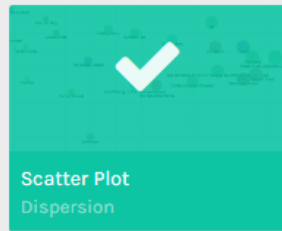
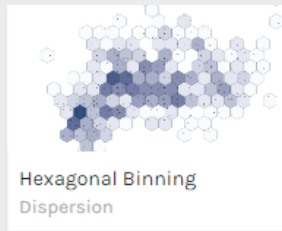
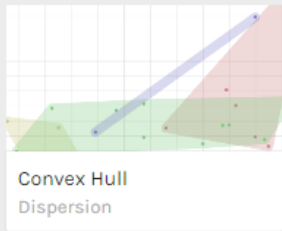
EXTERNAL LINK
[Project page](#)

Scatter Plot

Dispersion



A scatter plot, scatterplot, or scattergraph is a type of mathematical diagram using Cartesian coordinates to display values for two variables for a set of data. The data is displayed as a collection of points, each having the value of one variable determining the position on the horizontal axis and the value of the other variable determining the position on the vertical axis. This kind of plot is also called a scatter chart, scattergram, scatter diagram, or scatter graph.





Colors for data scientists. Generate and refine palettes of optimally distinct colors.

Color space

Default preset

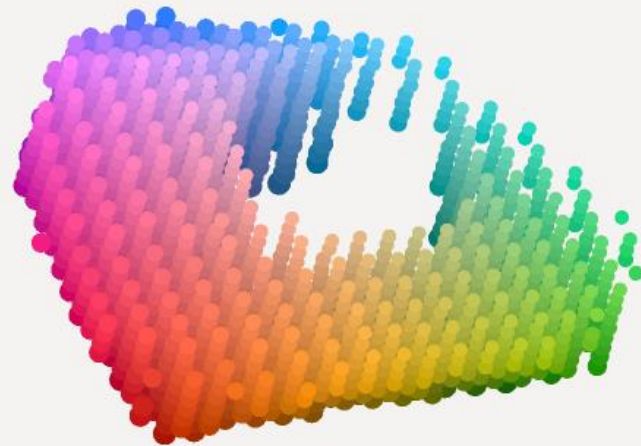
H 0 | 360

C 30 | 80

L 35 | 80

Improve for the **colorblind** (slow)

Dark background



Palette

5 colors | soft (k-Means)

Make a palette

5 color swatches

Tweet

We used: [Sigma.js](#), [Prettify](#), [Bootstrap](#), [jQuery](#), [Modernizr](#), [Initializr](#)
Check our [GitHub](#).

See also our other tools at [Médialab Tools!](#)

And a huge **thanks** to these inspiring works:

Chroma.js

I massively use this excellent js library to convert colors. If you have not done it yet, look at [this post](#). You'll understand much useful things about color in dataviz.

ColorBrewer

Very famous tool, that showed the way few years ago. If you do not know it, you *must* take a look.

SciencesPo | médialab

Developed by Mathieu Jacomy at the [Sciences-Po Medialab](#)

Help, bug report or contacting us: [GitHub Issues](#).

Nástroje pro vizualizaci dat (typologie)

- běžné kancelářské programy
 - tabulkové procesory apod., také online
- vizualizační nástroje s grafickým rozhraním (často i další funkce)
 - pro jeden typ vizualizace
 - komplexní
- moduly pro programovací jazyky
 - vizualizace často ve spojení s knihovnamí ve skriptovacím jazyce JavaScript
 - pro vytváření programů v daném jazyce slouží vývojové prostředí (IDE)
 - nejrozšířenější: R, Python

Nástroje pro vizualizaci dat (příklady)

- (jazyk) R + moduly
 - vývojové prostředí RStudio (<https://www.rstudio.com>) a další
- (jazyk) Python + moduly
 - vývojové prostředí Anaconda (<https://www.anaconda.com>) s integrovanými nástroji pro analýzu dat (zahrnuje i RStudio) a další



DOWNLOAD RSTUDIO

RStudio Desktop

Find out more about RStudio Desktop and RStudio Desktop Pro below.

DOWNLOAD RSTUDIO



RStudio interface showing R code execution and a map visualization.

```
1 library(leaflet)
2
3 this.dir <- dirname(parent.frame(2)$ofile)
4 setwd(this.dir)
5
6 lonlat = read.csv("konzistor.csv", sep = ";")
7
8
9 leaflet(data = lonlat) %>% addTiles(group = 'OSM') %>%
10
11 addMarkers(~longitude, ~latitude, popup=~id, label = ~as.character(id))
```

Environment History Connections

Global Environment

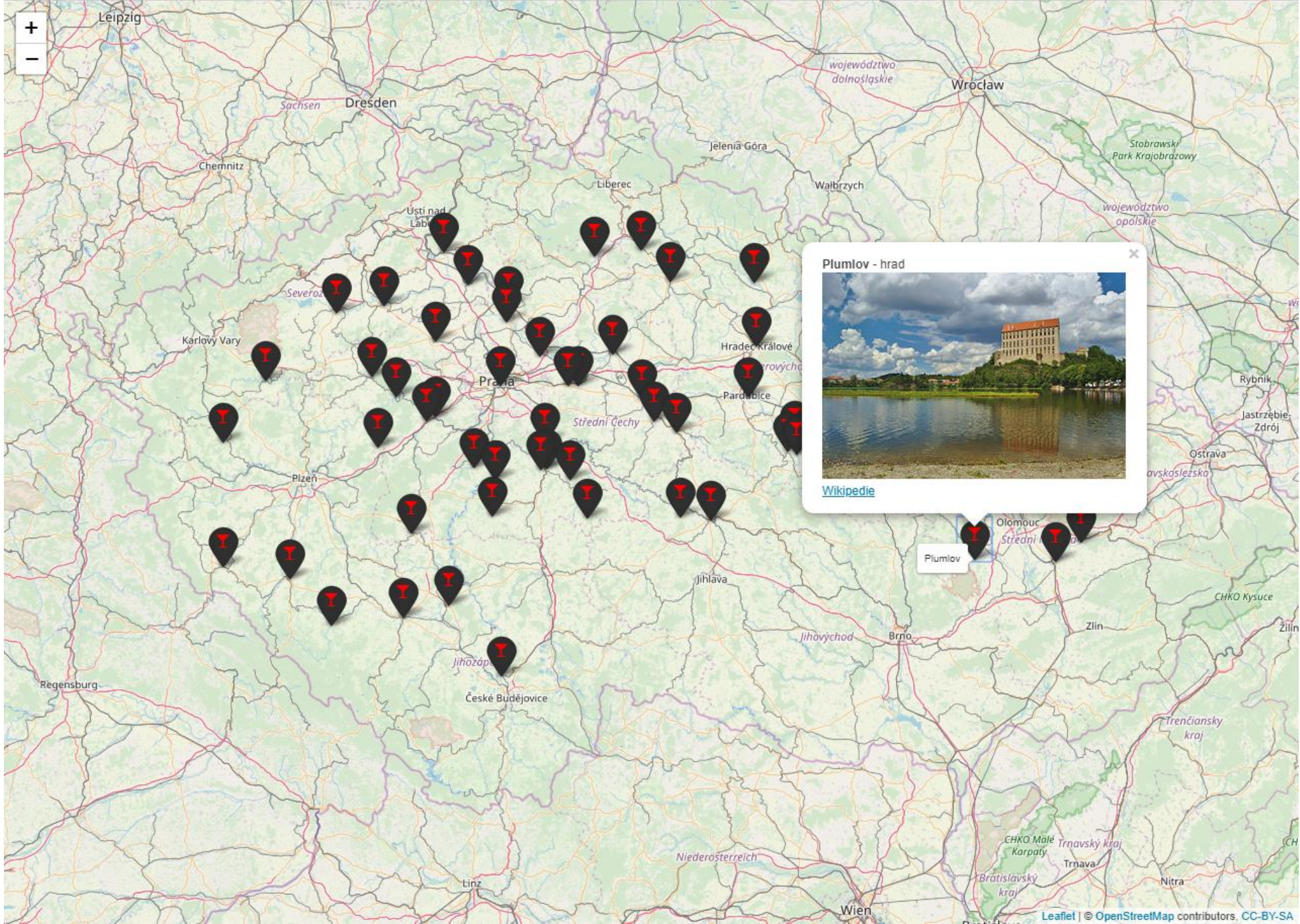
Data	
lonlat	52 obs. of 7 variables
values	
this.dir	"C:/Users/MAREKJ/OneDrive - Filozofická fakulta.."

Files Plots Packages Help Viewer

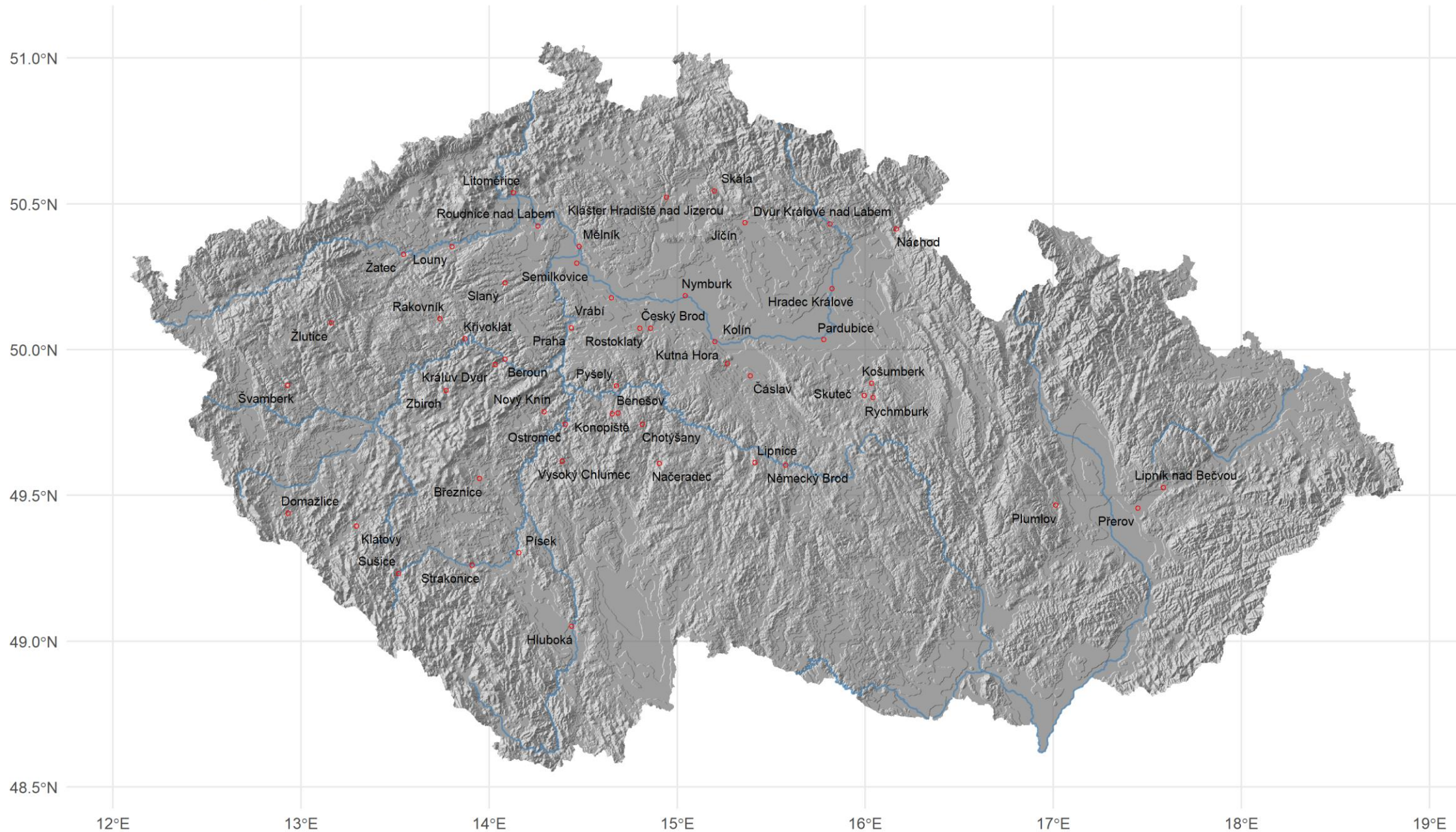
Console Terminal

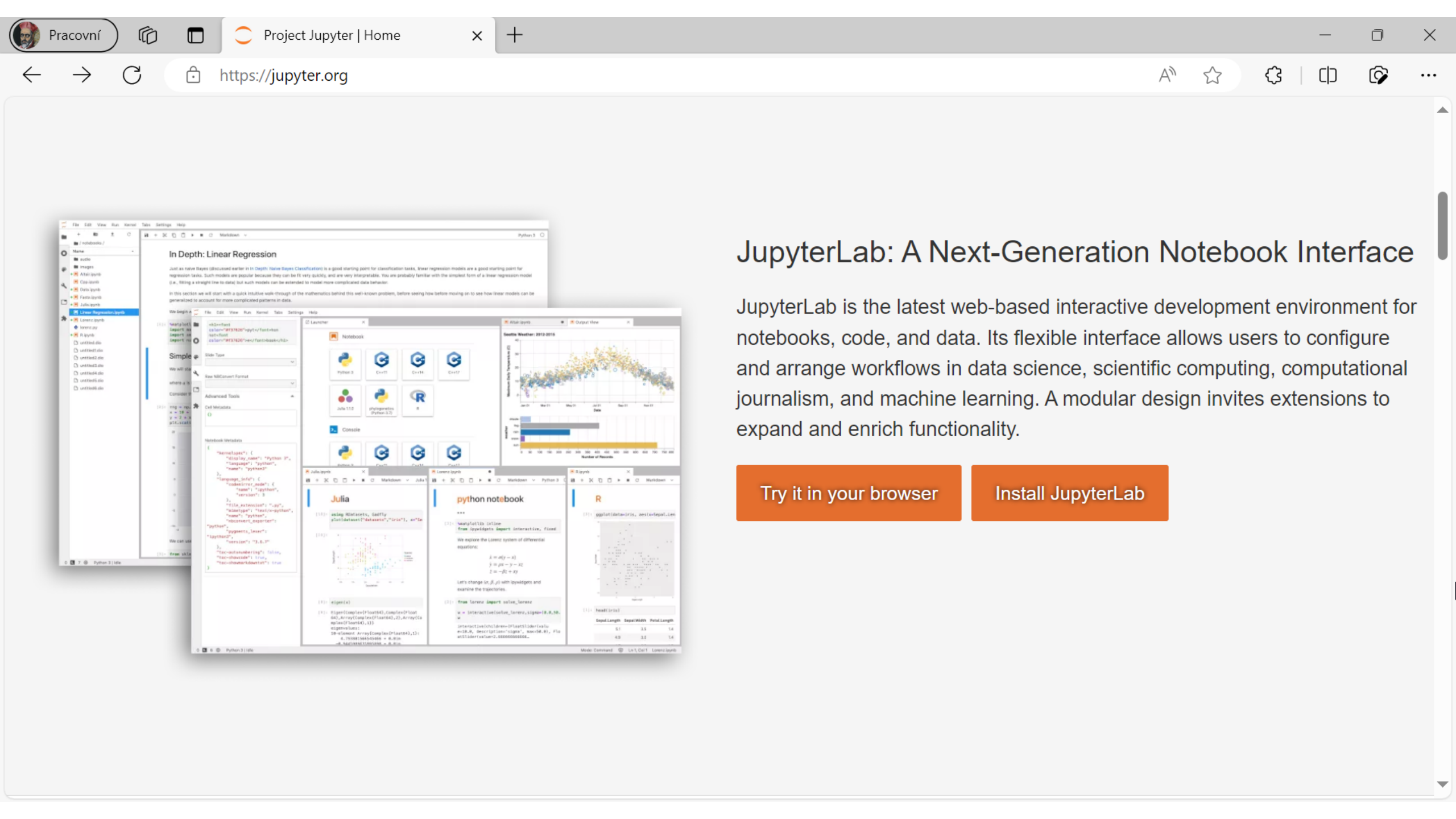
```
C:/Users/MAREKJ/OneDrive - Filozofická fakulta, Univerzita Karlova v Praze/Vyuka/proseminar_akademicke_prace_2018/vizualiza
> library(leaflet)
> this.dir <- dirname(parent.frame(2)$ofile)
> setwd(this.dir)
> lonlat = read.csv("konzistor.csv", sep = ";")
> leaflet(data = lonlat) %>% addTiles(group = 'OSM') %>%
+ addMarkers(~longitude, ~latitude, popup=~id, label = ~as.character(id))
>
>
>
>
>
>
>
>
>
>
>
```

Leaflet | © OpenStreetMap contributors, CC-BY-SA



Místa odeslání a přijetí konzistorních písemností (1471-1489)





JupyterLab: A Next-Generation Notebook Interface

JupyterLab is the latest web-based interactive development environment for notebooks, code, and data. Its flexible interface allows users to configure and arrange workflows in data science, scientific computing, computational journalism, and machine learning. A modular design invites extensions to expand and enrich functionality.

[Try it in your browser](#)

[Install JupyterLab](#)

```
In [1]: import folium
import pandas as pd
import folium.plugins

df = pd.read_csv('konzistor2.csv', delimiter=';')

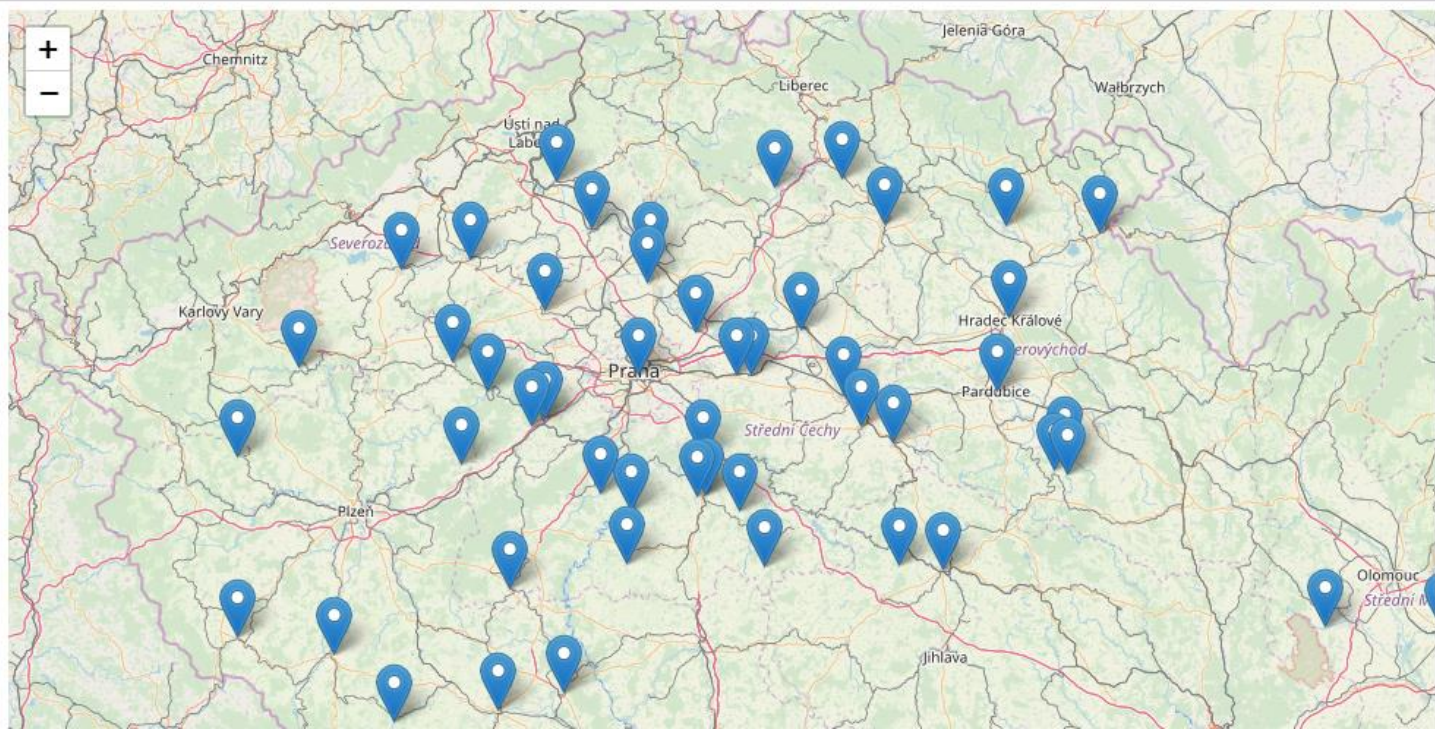
map_center = [df["lat"].mean(), df["lon"].mean()]
map_1 = folium.Map(location=map_center, zoom_start=8)

for i, row in df[["lat", "lon", "id"]].dropna().iterrows():
    position = (row["lat"], row["lon"])
    popup = (row["id"])
    folium.Marker(position, popup=popup).add_to(map_1)

map_1

display(map_1)

map_1.save('konzistor_python.html')
```



Nástroje s grafickým rozhraním

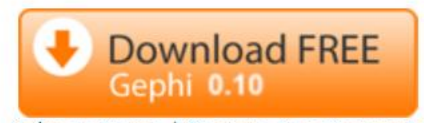
- jednoúčelové
- příklady
- <https://gephi.org/>



The Open Graph Viz Platform

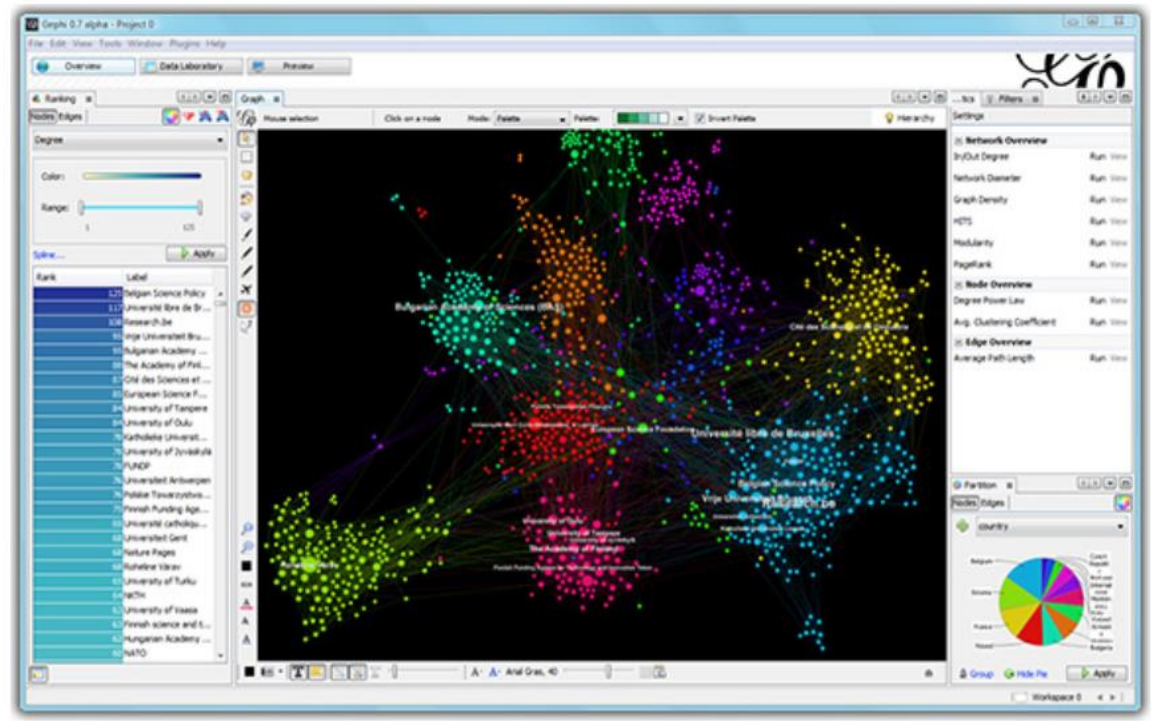
Gephi is the leading visualization and exploration software for all kinds of graphs and networks. Gephi is open-source and free.

Runs on Windows, Mac OS X and Linux.
[Learn More on Gephi Platform »](#)

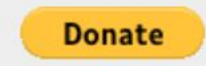


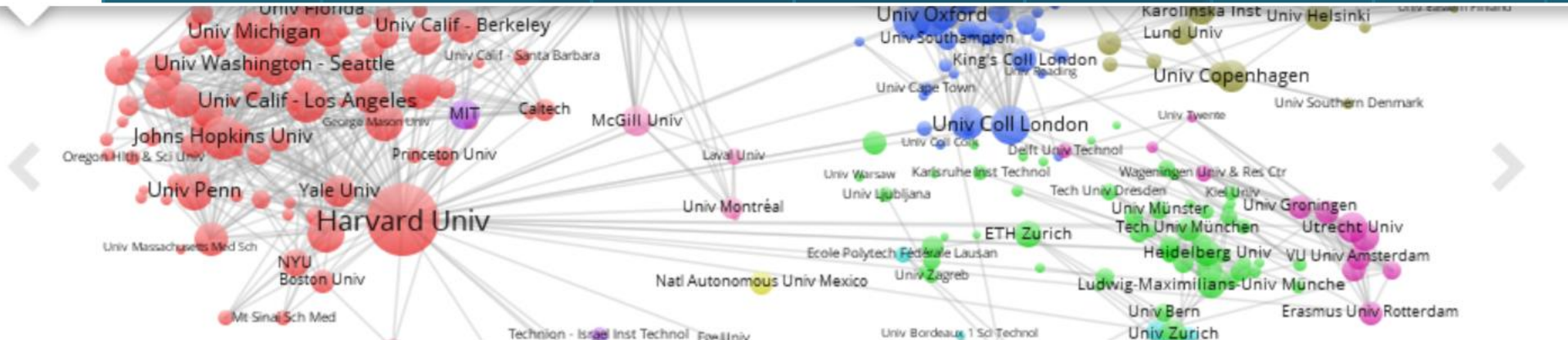
[Release Notes](#) | [System Requirements](#)

- ▶ [Features](#)
- ▶ [Quick start](#)
- ▶ [Screenshots](#)
- ▶ [Videos](#)



Support us! We are non-profit. Help us to **innovate** and **empower** the community by donating only 8€:





Welcome to VOSviewer

VOSviewer is a software tool for constructing and visualizing bibliometric networks. These networks may for instance include journals, researchers, or individual publications, and they can be constructed based on citation, bibliographic coupling, co-citation, or co-authorship relations. VOSviewer also offers text mining functionality that can be used to construct and visualize co-occurrence networks of important terms extracted from a body of scientific literature.

Nástroje s grafickým rozhraním

- komplexní
- příklady
- <http://heuristnetwork.org>
- <https://public.tableau.com>
- <https://powerbi.microsoft.com/en-us/desktop/>

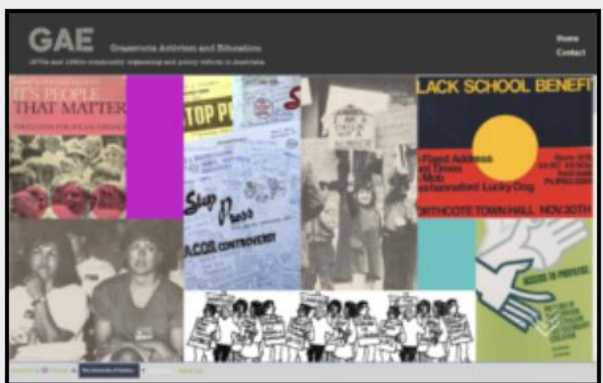
History ↑



Digital Harlem

An award-winning component of the Black Metropolis-ARC: Harlem collaborative research project, whose goal is to “produce an ethnographic study of everyday life in Harlem as it became the black capital of the world.” Since 2013, the back-end Heurist database is accessible via the project’s public website, to provide a research tool that gives a customised view of activities, places, and relationships relating to Harlem from 1915-1930. [more...](#)

Stephen Robertson (George Mason University)



Grassroots Activism in Education

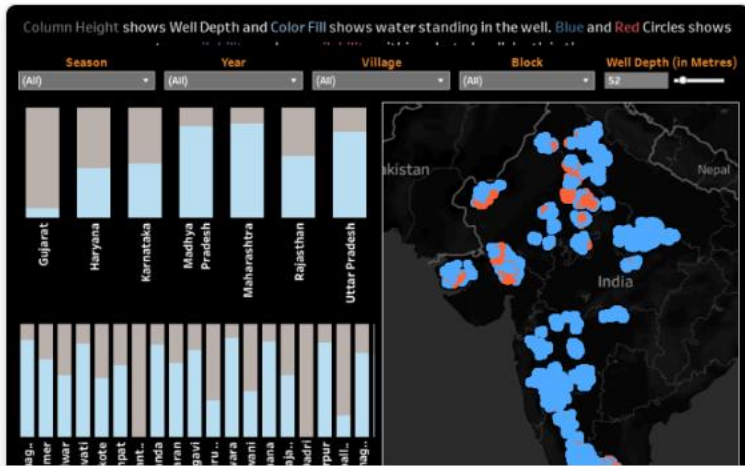
GAE is a website that highlights the work of ‘ **Community organising in Australian education policy, 1970s-1980s** ’, a three-year (2020-2023) research project investigating the phenomenon of ‘grassroots’ organising for education reform during a significant period in Australia’s history. The project is co-led by Professor Helen Proctor (University of Sydney), Associate Professor Jessica Gerrard (University of Melbourne) and Professor Susan Goodwin (University of Sydney) and funded by the Australian Research Council ([DP 200102378](#)).

The project has a strong visual identity, which structures their Heurist website. In the underlying

Trending

See All

Explore vizzes that are popular right now



India Ground Water Level Trends

Amrit Kumar

☆ 66 👁 2,378

10 HACKS

to bring your dashboard from good to GREAT

- Different scales
- Different formats
- Static Titles
- Extended Axis
- Different Color
- Special Characters
- Scalable Legends
- Using Containers
- SVG vs SVG
- Show everything

10 hacks to make your dashboard GREAT!

Louis Yu

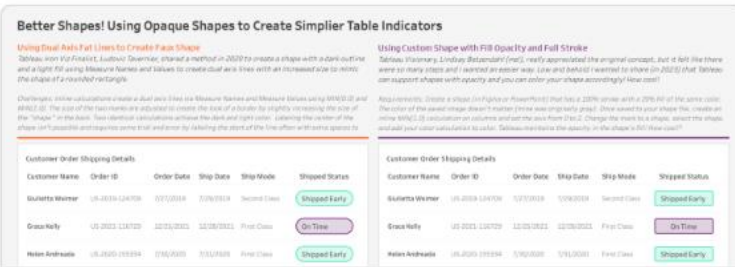
☆ 333 👁 6,234

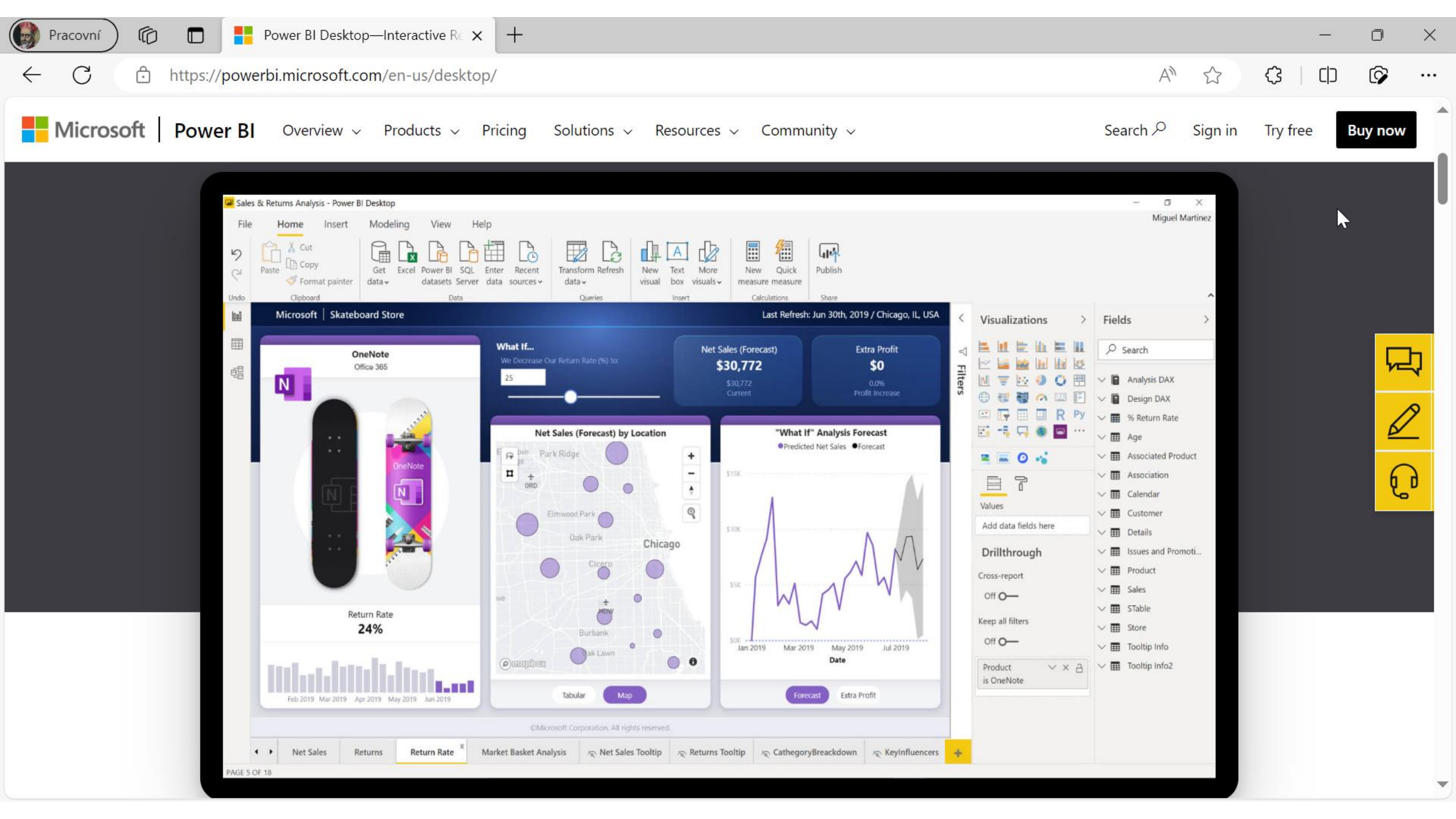


Roman Empire Establishment 79BC-117AD

Nir Smilga

☆ 270 👁 12,633





Pracovnf



Power BI Desktop—Interactive R



https://powerbi.microsoft.com/en-us/desktop/



Power BI

Overview

Products

Pricing

Solutions

Resources

Community

Search

Sign in

Try free

Buy now

Sales & Returns Analysis - Power BI Desktop

Miguel Martinez

File Home Insert Modeling View Help

Undo

Paste Copy Format painter

Clipboard

Get data Excel Power BI datasets SQL Server Enter data Recent sources

Data

Transform data Refresh data

Queries

New visual Text box More visuals

Insert

New Quick measure measure

Calculations

Publish

Share

Microsoft | Skateboard Store

Last Refresh: Jun 30th, 2019 / Chicago, IL, USA

OneNote Office 365

Return Rate **24%**

Feb 2019 Mar 2019 Apr 2019 May 2019 Jun 2019

What If...

We Decrease Our Return Rate (%) to:

25

Net Sales (Forecast)

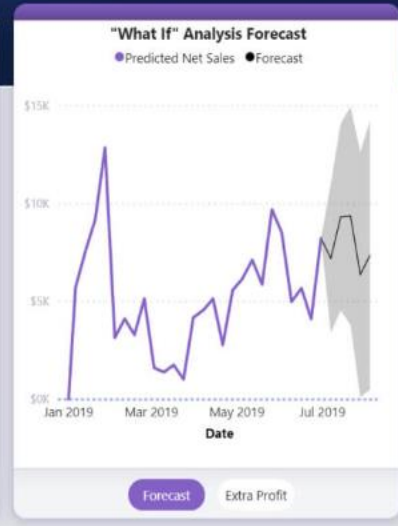
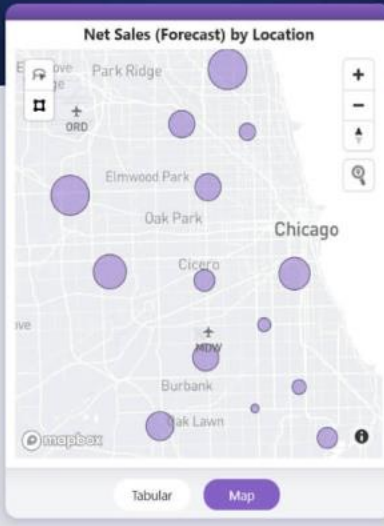
\$30,772

\$30,772 Current

Extra Profit

\$0

0.0% Profit Increase



Filters

Visualizations

Values

Add data fields here

Drillthrough

Cross-report

Off

Keep all filters

Off

Product is OneNote

Fields

Search

- Analysis DAX
- Design DAX
- % Return Rate
- Age
- Associated Product
- Association
- Calendar
- Customer
- Details
- Issues and Promoti...
- Product
- Sales
- STable
- Store
- Tooltip info
- Tooltip info2

Literatura

- CHEN, Chun-houh, ed., HÄRDLE, Wolfgang, ed. a UNWIN, Antony, ed. *Handbook of data visualization*. Berlin: Springer, 2008. xiii, 936 s. Springer handbooks of computational statistics. ISBN 978-3-540-33036-3.
- *Tools & Tutorials* (University of Toronto Map & Data Library. Dostupné z www: <https://mdl.library.utoronto.ca/dataviz/tools-tutorials> [cit. 15. 11. 2018].

Děkuji vám za pozornost

Kontakt: jindrich.marek@ff.cuni.cz