|  |  |  |
| --- | --- | --- |
| **Univerzita Hradec Králové**  **Fakulta informatiky a managementu**  **Název katedry** | | |
| **Název bakalářské/diplomové práce**  (podtitul práce)  Bakalářská/Diplomová práce | | |
| Autor: Jméno, příjmení  Studijní obor: Studijní obor | | |
| Vedoucí práce: Titul, jméno, příjmení  Odborný konzultant: Titul, jméno, příjmení  Pracoviště | | |
| Hradec Králové | měsíc rok | |
| Prohlášení:  Prohlašuji, že jsem bakalářskou/diplomovou práci zpracoval/zpracovala samostatně a s použitím uvedené literatury. | | | |
| V Hradci Králové dne 19.9.2021 | | *vlastnoruční podpis*  Jméno a Příjmení | |
| Poděkování:  Děkuji vedoucímu bakalářské/diplomové práce titul, jméno, příjmení za metodické vedení práce a…. | | | |

Anotace

Text anotace – shrnutí cíle, významu práce a výsledky v ní dosažené. Délka minimálně 100 a maximálně 200 slov.

Annotation

Title: Název práce v anglickém jazyce

Anotace v anglickém jazyce. Délka minimálně 100 a maximálně 200 slov.

Obsah

[1 Úvod 1](#_Toc82354828)

[2 Cíl práce 2](#_Toc82354829)

[3 Metodika zpracování 3](#_Toc82354830)

[4 Teoretická část 4](#_Toc82354831)

[4.1 Jazyk Python 4](#_Toc82354832)

[4.2 Vizualizace 4](#_Toc82354833)

[4.2.1 Vizualizace obecně 4](#_Toc82354834)

[4.2.2 Stručná historie vizualizace dat 4](#_Toc82354835)

[4.2.3 Barvy v datové vizualizaci 4](#_Toc82354836)

[4.3 Vizualizační knihovny jazyka Python 4](#_Toc82354837)

[4.3.1 Matplotlib 4](#_Toc82354838)

[4.3.2 Seaborn 4](#_Toc82354839)

[4.3.3 Bokeh 4](#_Toc82354840)

[4.3.4 Plotly 4](#_Toc82354841)

[4.3.5 Pygal 4](#_Toc82354842)

[4.3.6 Holoviews 4](#_Toc82354843)

[4.3.7 MidiTime 4](#_Toc82354844)

[4.3.8 Geoplotlib 4](#_Toc82354845)

[4.3.9 WordCloud 4](#_Toc82354846)

[4.4 Způsoby získávání dat 4](#_Toc82354847)

[4.4.1 4](#_Toc82354848)

[4.5 Způsoby zpracování dat? 4](#_Toc82354849)

[5 Praktická část 5](#_Toc82354850)

[5.1 Vytvořené vizualizace 5](#_Toc82354851)

[5.2 STAG projekt? 5](#_Toc82354852)

[5.2.1 Podřazená podkapitola 6](#_Toc82354853)

[6 Shrnutí výsledků 7](#_Toc82354854)

[7 Závěry a doporučení 8](#_Toc82354855)

[8 Seznam použité literatury 9](#_Toc82354856)

[9 Přílohy 10](#_Toc82354857)

Seznam obrázků

[Obr. 1 Název obrázku/grafu/fotografie. 3](#_Toc348517265)

Seznam tabulek

[Tabulka 1 Název tabulky. 3](#_Toc348517268)

# Úvod

Zde vysvětlit problémovou situaci a otázky, které se budou v bakalářské/diplomové práci řešit.

# Cíl práce

Smysl a účel, výzkumné otázky.

# Metodika zpracování

Cíle, hypotézy/ výzkumné otázky, způsob hledání odpovědí na výzkumné otázky včetně metodiky vlastního výzkumu/šetření, literární rešerše.

# Teoretická část

## Jazyk Python

## Vizualizace

### Vizualizace obecně

Definice pojmu „vizualizace dat/informací“ není až tak jednoznačná a jednoduchá, jak by se na první pohled mohlo zdát.

Kniha „Lecture Notes in Computer Science“ z roku 1970 popisuje, že „*Vizualizace informací využívá počítačové grafiky a interakce, aby napomohla lidem v řešení problémů*“ [1]. Vizualizace však nemusí nutně sloužit pouze k tomuto účelu, může najít uplatnění například ve výuce matematiky, kde „*Matematické koncepty, jako čísla, funkce, nebo vektory […] nejsou součástí intuitivního chápání, tak jako reálné fyzické objekty*“ [2]. Vizualizací je však možné pochopení těchto konceptů studentům usnadnit. „*Informační vizualizace nabízí možnost snáze a lépe ukázat klasické vizuální reprezentace matematických formátů, ale také je obohatit o prvky pohybu a interaktivity […] informační vizualizace tedy plní didaktické funkce nezbytné pro výuku matematiky.*“ [2]

Vizualizace však může plnit i dekorativní roli, Lorène Fauvelle v článku „Data visualization: definition, examples, tools, advice“ popisuje „Datové umění“ jako nejvyšší formu vizualizace kde „*Získávání informací a postřehů již nestačí, výsledek musí být také vizuálně atraktivní…* “ [3] v rozhovoru s Nicholasem Rougeux popisuje, že „*Datový umělec se nesnaží pouze informovat, ale hlavně vyvolat emoce*“. [4]

Většina zdrojů se však shoduje na dvou základních konceptech „redukce“ a „prostoru“, které popisuje i Lev Manovich v článku „What is Visualization?“: „*Vizualizace informací využívá jednoduchých grafických elementů jako zástupce reálných objektů a vazeb mezi nimi – nezáleží na tom, jestli se jedná o lidi, ceny na burze, příjmy států, nezaměstnanost, nebo cokoliv jiného. Skrze tyto grafické elementy informační vizualizace odhaluje vzorce a struktury v datech. Nicméně, cenou za tuto možnost je extrémní úroveň schematizace. Zahazujeme 99 % toho, co je na objektech specifické, ve snaze nalezení vzorců na zbývajícím 1 % charakteristik.*“ [7]

„*Co mají všechny vizualizační techniky společné, kromě redukce? Všechny využívají prostoru (umístění, velikosti, tvaru a v poslední době i zakřivení a pohybu) pro reprezentaci klíčových rozdílů v datech a zobrazení nejdůležitějších vzorců a vztahů. […] Ostatní, méně důležité vlastnosti objektů jsou pak reprezentovány jinými vizuálními prostředky – odstínem, barvou, vzorem stínování nebo i průhledností.* “ [7]

### Stručná historie vizualizace dat

### Barvy v datové vizualizaci

## Vizualizační knihovny jazyka Python

### Matplotlib

Pravděpodobně nejznámější vizualizační knihovna jazyka Python byla vytvořena okolo roku 2003 Johnem Huterem byla původně určena k vizualizaci dat zaznamenaných elektrokortikografií při výzkumu epilepsie. Laboratoř, ve které John Hunter pracoval, měla v té době pouze jednu licenci na softwarový balíček pro analýzu dat, o který se museli všichni výzkumníci dělit. Hunter se tedy rozhodl vytvořit náhradu v prostředí MATLAB. Tato aplikace však nebyla ideální pro vizualizaci dat z mnoha zdrojů (kromě elektrokortikografie například i EEG a magnetické rezonance) uložených na několika serverech. Hunter proto začal vyvíjet novou aplikaci v jazyce Python „EEG viewer and analyzer“, která se postupem času změnila v dnešní knihovnu matplotlib. [5]

V současné době se jedná o volně dostupný open-source projekt, sloužící k tvorbě statických, animovaných i interaktivních vizualizací nejčastěji ve 2D (knihovna však podporuje i trojrozměrné vizualizace). [6]

### Seaborn

Seaborn vznikl jako řešení nejčastějších nedostatků knihovny matplotlib, konkrétně defaultních nastavení vizualizace před verzí 2.0, která byla založena na vizualizacích MATLABu, nižší úrovně API u matplotlibu, která často měla za následek nadbytečný kód a složitost využití datových struktur knihovny pandas. [9]

Seaborn buduje na základech knihovny matplotlib, ke které poskytuje vysokoúrovňové rozhraní, také dokáže využívat datové struktury knihovny pandas. Seaborn na základě specifikace typu grafu zvládne automaticky provázat hodnoty v datech s vizuálními atributy, jako je barva, velikost a styl, propočítat statistické transformace a doplnit ke grafu informativní štítky a legendu. Díky tomu, že Seaborn dokáže vytvořit kompletní vizualizace jedním voláním funkce s minimálním počtem argumentů, je ideálním nástrojem pro explorační analýzu dat. [8]

### Bokeh

Stejně jako matplotlib je knihovna Bokeh open source projektem, finančně podporovaným neziskovou organizací NumFOCUS, která se zaměřuje na financování a propagaci, open source nástrojů ve vědě a výzkumu. [10]

Oproti knihovnám matplotlib a seaborn však Bokeh nevytváří kompletní grafy jednou funkcí, místo toho nabízí širokou škálu nástrojů pro manipulaci s elementy vektorové grafiky (nazývanými „glyphs“), ze kterých lze následně vytvářet jednotlivé vrstvy vizualizace. To dává uživateli mnohem větší kontrolu nad vzhledem finální vizualizace, za cenu větší složitosti kódu, potřebného na její vytvoření. Bokeh dále dovoluje do vizualizace přidávat interaktivní elementy, jako slidery, tlačítka nebo drop-down menu. Knihovna také umožňuje vytvoření „Bokeh serveru“, který zajišťuje možnost streamování dat a složitější uživatelské interakce.

Knihovna se vlastně skládá ze dvou, Bokeh pro Python – sloužící k vytváření vizualizací a BokehJS – knihovna jazyka JavaScript sloužící k renderování vizualizace a zajištění interaktivity ve webovém prohlížeči. [11]

### Plotly

Další open source knihovnou pro datovou vizualizaci je Plotly. Tato knihovna však není omezena pouze na jeden jazyk, lze ji používat v jazycích Python, R, Julia, MATLAB a existují i projekty zpřístupňující ji v jazyce Java a jazycích frameworku .NET. Implementace v jazyce Python je však její zdaleka nejoblíbenější variantou a existují zde dvě možnosti využití knihovny. Jednodušší vysokoúrovňový Plotly Express, který s minimem kódu vytváří kompletní vizualizace a složitější modul Plotly Graphic Objects, který však dovoluje mnohem větší kontrolu nad vytvářenou vizualizací a je pomocí něho možné vytvořit některé typy grafů, nepodporované modulem Express. [12]

Existují také komerční produkty založené na této knihovně, Dash Enterprise - platforma pro analýzu trhu, datové vědy a výzkum v oblasti umělé inteligence a Chart Studio Enterprise – nástroj pro rychlou tvorbu vizualizací přímo ze souborů nebo databází a jejich vkládání do webových stránek. [12]

### Holoviews

### Pygal

### MidiTime

### Geoplotlib

### WordCloud

## Způsoby získávání dat

## Způsoby zpracování dat?

# Praktická část

## Vytvořené vizualizace

## STAG projekt?

Vlastní text práce.

Tabulka Název tabulky.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Zdroj: citace zdroje, nebo autor, vlastní zpracování

### Podřazená podkapitola

Vlastní text práce.



Obr. Název obrázku/grafu/fotografie.

Zdroj: citace zdroje, nebo autor, vlastní zpracování

# Shrnutí výsledků

Souhrn vlastních výsledků získaných v průběhu řešení problému.

# Závěry a doporučení

Kritická diskuze nad výsledky, ke kterým autor dospěl (soulad výsled-ků  literaturou či předpoklady; výsledky a okolnosti, které zvláště ovlivnily předkládanou práci atd.). Je vhodné naznačit i případné další (popř. alternativní) možnosti zkoumání dané problematiky a otevřené problémy pro další studium.

# Seznam použité literatury

1. PURCHASE, Helen et al. Theoretical Foundations of Information Visualization. In: *Lecture Notes In Computer Science*. 4950. 1970, s. 46–64. ISBN 978-3-540-70955-8. DOI: [10.1007/978-3-540-70956-5\_3](https://doi.org/10.1007/978-3-540-70956-5_3)
2. G. CHIAPPINI a R.M. BOTTINO. *Visualisation in Teaching-Learning Mathematics: The Role of the Computer* [online]. Dostupné z: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.39.3360&rep=rep1&type=pdf>
3. LORÈNE FAUVELLE. *Data visualization: definition, examples, tools, advice [guide 2021]* [online]. 2020 [cit. 03.04.2021]. Dostupné z: <https://www.intotheminds.com/blog/en/data-visualization/>
4. INTOTHEMINDS. *What is a data artist? | with Nicholas Rougeux* [online]. 2020 [cit. 16.09.2021]. Dostupné z: <https://www.youtube.com/watch?v=k4D9qgVb17Q>
5. JOHN HUNTER a MICHAEL DROETTBOOM. The Architecture of Open Source Applications (Volume 2): matplotlib. In: [cit. 16.09.2021]. Dostupné z: <http://aosabook.org/en/matplotlib.html>
6. JOHN HUNTER a DARREN DALE. Overview — Matplotlib 3.4.3 documentation. In: . 2021 [cit. 08.09.2021]. Dostupné z: <https://matplotlib.org/stable/contents.html>
7. MANOVICH, Lev. What is visualisation? *Visual Studies*. Routledge, 2011, roč. 26, č. 1, s. 36–49. ISSN 1472-586X. DOI: [10.1080/1472586X.2011.548488](https://doi.org/10.1080/1472586X.2011.548488)
8. WASKOM, Michael. seaborn: statistical data visualization. *Journal of Open Source Software*. 2021, roč. 6, č. 60, s. 3021. ISSN 2475-9066. DOI: [10.21105/joss.03021](https://doi.org/10.21105/joss.03021)
9. VANDERPLAS, Jake. *Python Data Science Handbook: Essential Tools for Working with Data*. 1st edition. vyd. Sebastopol, CA: O’Reilly Media, 2016. ISBN 978-1-4919-1205-8.
10. Sponsored Projects | pandas, NumPy, Matplotlib, Jupyter, + more. In: *NumFOCUS* [online] [cit. 19.09.2021]. Dostupné z: <https://numfocus.org/sponsored-projects>
11. BOKEH CONTRIBUTORS. Bokeh documentation. 2021 [cit. 08.09.2021]. Dostupné z: <https://docs.bokeh.org/en/latest/index.html>
12. Plotly Open Source Graphing Libraries. [cit. 19.09.2021]. Dostupné z: <https://plotly.com/api/>
13. A
14. A
15. A
16. A
17. A
18. a
19. a

# Přílohy

*Oskenované zadání práce*