



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

FACULTY OF INFORMATION TECHNOLOGY

ÚSTAV INFORMAČNÍCH SYSTÉMŮ

DEPARTMENT OF INFORMATION SYSTEMS

VIZUALIZACE STATISTICKÝCH DAT ČESKÉ REPUBLIKY

VISUALIZATION OF CZECH STATISTICAL DATA

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

HANA RYBNIKÁŘOVÁ

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. JIŘÍ HYNEK

BRNO 2019

Zadání bakalářské práce



21705

Studentka: **Rybníkářová Hana**
Program: Informační technologie
Název: **Vizualizace statistických dat České republiky**
Visualization of Czech Statistical Data
Kategorie: Informační systémy

Zadání:

1. Seznamte se s principy tvorby informačních systémů s důrazem na vizualizaci dat a nástroj *dashboard*.
2. Analyzujte dostupné databáze poskytující statistická data České republiky (demografická, ekonomická, sociální apod.). Analyzujte existující aplikace a informační systémy určené pro daný účel. Zamyslete se nad novými a efektivními způsoby prezentace vybraných statistických dat. Konzultujte tyto způsoby s vedoucím práce.
3. Navrhněte způsob získávání a ukládání vybraných datových sad z bodu 2.
4. Navrhněte informační systém, který bude vhodně vizualizovat vybraná statistická data. Klad'te důraz na vizualizační principy a srozumitelnost dat.
5. Navržený informační systém implementujte.
6. Otestujte funkcionální a použitelnost výsledné systému. Navrhněte rozšíření o další datové sady.

Literatura:

1. Few, S.: *Information Dashboard Design: The Effective Visual Communication of Data*. Sebastopol, USA: O'Reilly, 2006, ISBN 978-059-6100-162.
2. Edward, T.: *The visual display of quantitative information*. Cheshire, USA: Graphics Press, 2001, ISBN 978-0-9613921-4-7.

Pro udělení zápočtu za první semestr je požadováno:

- Body 1 až 4.

Podrobné závazné pokyny pro vypracování práce viz <http://www.fit.vutbr.cz/info/szz/>

Vedoucí práce: **Hynek Jiří, Ing.**
Vedoucí ústavu: Kolář Dušan, doc. Dr. Ing.
Datum zadání: 1. listopadu 2018
Datum odevzdání: 15. května 2019
Datum schválení: 25. října 2018

Abstrakt

Tato bakalářská práce popisuje návrh a implementaci informačního systému, který vizualizuje statistická data České republiky z oblasti zaměstnanosti, nezaměstnanosti a průměrné mzdy obyvatel. Jeho cílem je přiblížení takových informací studentům, kteří se například rozhodují, jaké povolání zvolit. Data jsou čerpána z Českého statistického úřadu a Ministerstva práce a sociálních věcí a pro vhodnou vizualizaci byl využit nástroj dashboard, který klade důraz na zobrazení všech dat na jednu obrazovku.

Abstract

This bachelor's thesis describes design and implementation of information system, which visualizes Czech statistical data of employment, unemployment and average salary. The goal is to approach these informations to students who need to determine their future job. Data are taken from Czech statistical office and Ministry of Labour and Social Affairs. For visualization was chosen information dashboard which is visual display of the most important information arranged on a single screen.

Klíčová slova

statistická data, zaměstnanost, nezaměstnanost, mzdy, studenti, vizualizace, dashboard, informační systém, grafy

Keywords

statistical data, employment, unemployment, salary, students, visualization, dashboard, information system, charts

Citace

RYBNIKÁŘOVÁ, Hana. *Vizualizace statistických dat České republiky*. Brno, 2019. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Ing. Jiří Hynek

Vizualizace statistických dat České republiky

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracovala samostatně pod vedením pana Ing. Jiřího Hynka a uvedla jsem všechny literární prameny a publikace, ze kterých jsem čerpala.

.....
Hana Rybníkářová
14. května 2019

Poděkování

Na tomto místě bych chtěla poděkovat vedoucímu práce Ing. Jiřímu Hynkovi za trpělivost a cenné rady. Dále bych ráda poděkovala svému příteli a celé rodině za velkou podporu.

Obsah

1	Úvod	3
2	Statistická data ČR	4
2.1	Český statistický úřad	4
2.2	Veřejná databáze ČSÚ	4
2.3	Kategorizace dat	6
2.4	Výběr dat	7
2.4.1	Zaměstnanost, nezaměstnanost	7
2.4.2	Průměrné mzdy	8
2.4.3	Vzdělávání	9
2.4.4	Kriminalita, nehody	9
2.5	Další zdroje dat	9
3	Dashboard a vizualizace dat	11
3.1	Definice dashboardu	11
3.2	Klasifikace dashboardů	12
3.3	Tvorba dashboardu	13
3.3.1	Identifikace uživatelů	13
3.3.2	Stanovení metrik	14
3.3.3	Vizualizace	15
3.3.4	Testování	17
4	Analýza uživatelů a požadavků	18
4.1	Průzkum uživatelů	18
4.2	Současná dostupnost dat	20
5	Návrh systému a získání dat	22
5.1	Stanovení metrik	22
5.2	Návrh obrazovek	23
5.3	Responzivita	24
5.4	Návrh databáze	25
6	Použité technologie	27
6.1	Python	27
6.2	MySQL	27
6.3	PHP	27
6.4	HTML, CSS, JavaScript	28

7 Implementace	30
7.1 Zpracování dat	30
7.2 Informační systém	30
7.2.1 Prvky obrazovek	31
7.2.2 Rozložení prvků	34
8 Testování	35
8.1 Průběžné testování	35
8.2 Test použitelnosti	35
8.3 Zhodnocení testů	36
9 Závěr	37
Literatura	38
A Obsah CD	40
B Další vytvořená persona	41
C Snímky obrazovek	42
D Zkratky	44

Kapitola 1

Úvod

Statistická data České republiky dnes zpracovává a poskytuje Český statistický úřad (ČSÚ). Na jeho stránkách jsou publikována data v různých verzích jako jsou online publikace, časové řady, analýzy, ročenky nebo časopisy. Pro vyhledání konkrétních dat byla vytvořena tzv. *Veřejná databáze*, kde se uživatel dostane téměř ke všem informacím. Zobrazení dat ve veřejné databázi je ale složité a pro uživatele těžce pochopitelné. Vhodná vizualizace dat je přitom velkým přínosem k rychlému pochopení dat.

V rámci této práce jsou rozebrány problémy aktuálního zobrazení statistických dat a je navržen informační systém, který se snaží těmto problémům vyvarovat. Oblastí, ze kterých ČSÚ data poskytuje, je mnoho, proto byla pro první návrh a vytvoření systému vybrána data z oblasti zaměstnanosti, průměrných platů, vzdělání a kriminality. Tato data se dají rozlišovat podle několika kritérií a mohou poskytnout zajímavé informace například pro studenty středních škol, které zajímá jejich budoucí možné povolání.

Vhodným nástrojem k vizualizaci daných dat je dashboard. Tento vizualizační nástroj klade důraz na zobrazení a rozvržení velkého množství dat tak, aby poskytované informace byly srozumitelné a lépe pochopitelné. Využívá k tomu převážně grafické prvky jako jsou různé typy grafů nebo zobrazení na mapě. Důležitým bodem je přehled všech dat na jedné obrazovce, aby uživatel nemusel informace složitě hledat.

Kapitola 2 se zabývá problematikou statistických dat, kde jsou získávána, publikována a jak jsou pro běžné uživatele dostupná. V kapitole 3 je popsán vizualizační nástroj dashboard, jeho klasifikace a tvorba. Kapitoly 4 a 5 obsahují průzkum uživatelů, analýzu současné dostupnosti dat a návrh informačního systému. Poslední část je věnována použitým technologiím, implementaci a testování (kapitoly 6, 7 a 8).

Kapitola 2

Statistická data ČR

Statistické informace představují agregované hodnoty, které vznikají sběrem dat ze sociální, demografické nebo ekonomické oblasti. Agregací se rozumí součet, průměr, medián, minimum nebo maximum. Tato data mohou být následně využívána na různé přehledy, analýzy a plánování. V České republice získávání a zpracování údajů pro statistické účely zabezpečuje Český statistický úřad [2].

Velkým přínosem k publikování a získávání dat je Veřejná databáze, která je v této kapitole probírána. Dále jsou blíže popsány oblasti dat, která ČSÚ poskytuje, a nakonec jsou podrobně charakterizována data, která byla vybrána pro tvorbu nového informačního systému.

2.1 Český statistický úřad

ČSÚ je ústředním orgánem státní správy ČR. Byl zřízen dne 8. ledna 1969 zákonem č. 2/1969 Sb., o zřízení ministerstev a jiných ústředních orgánů státní správy. Vedle získávání a zpracování údajů poskytuje státním orgánům, orgánům územní samosprávy, veřejnosti a do zahraničí statistické informace o sociálním, ekonomickém, demografickém a ekologickém vývoji České republiky, také vede registr ekonomických subjektů, registr sčítacích obvodů a budov a zemědělský registr, a vykonává státní správu na úseku zpracování výsledků voleb do zastupitelstev územních samosprávních celků nebo Parlamentu České republiky [1]. Zajišťuje vzájemnou srovnatelnost statistických informací ve vnitrostátním i mezinárodním měřítku a jednou za deset let probíhá sčítání lidu, domů a bytů.

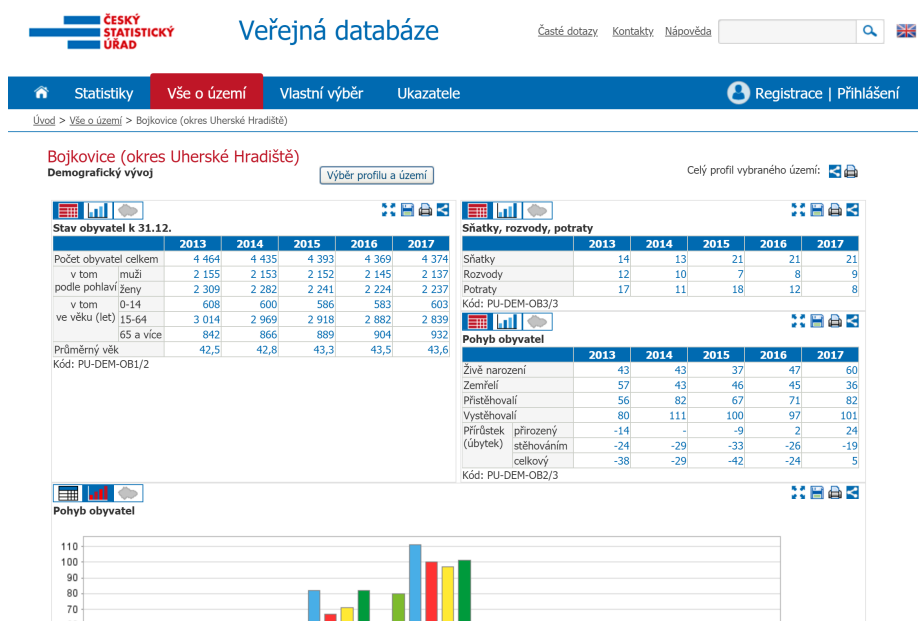
Data a informace jsou publikovány ve Statistické ročence České republiky a v dalších odborných publikacích a časopisech. ČSÚ zajišťuje i vývoj nových statistických metod. Veškerá data a informace jsou dostupné online¹. Najdeme zde různé analýzy, časové řady, datové sady nebo online publikace. V rámci tzv. *Rychlých informací* jsou zveřejňovány nejnovější údaje o vývoji ČR. V roce 2006 vznikla pro zobrazování dat i tzv. *Veřejná databáze*, která je podrobně popsána v další části.

2.2 Veřejná databáze ČSÚ

Veřejná databáze představuje ucelený informační systém, který je určen pro širokou laickou i odbornou veřejnost. Jsou zde k dispozici statistické údaje ze všech oblastí, které jsou zmíněny v další části. Uživatel má možnost projít si všechny tabulky, ale také se může

¹<http://www.czso.cz>

dostat do sekce „Vše o území“, která nabízí demografický přehled o dané oblasti v podobě několika grafů a tabulek najednou. Zobrazení a grafická prezentace jsou ale zastaralé a pro náročnější uživatele téměř nepoužitelné.



Obrázek 2.1: Příklad zobrazení informací o území ve Veřejné databázi

V neposlední řadě Veřejná databáze také nabízí možnost vlastního výběru. Zde si uživatel může zvolit ukazatele z různých oblastí a zobrazit je pro dané území a období. Přihlášení uživatelé si mohou výběry i ukládat. Důležitým bodem ale je, že všechny tabulky a vytvořené výběry se dají stáhnout a to v několika formátech. Patří mezi ně XLSX, XML a PDF.

Standardním prvkem všech údajů ve veřejné databázi jsou metainformace. Jednotlivé ukazatele používané ve statistice přesně definuje statistický metainformační systém [20]. V něm jsou obsažena statistická metadata, která zahrnují popisy objektů. Poskytují tak informaci o datech, o procesech tvorby a užití dat. Metainformace jsou rozděleny na klasifikace, číselníky a ukazatele.

Klasifikace představují hierarchické uspořádání určitých jevů či procesů, které jsou zpravidla uspořádány do tříd, podtříd, skupin a podskupin. Statistické klasifikace se vyznačují důkladným pokrytím reality, vzájemně se vylučujícími kategoriemi a metodologickými principy. Většina dnes používaných klasifikací byla vytvořena podle standardů Evropské unie, aby byla data srovnatelná se všemi zeměmi EU.

V této práci budou nejvíce využívány klasifikace zaměstnání (CZ-ISCO) a klasifikace ekonomických činností (CZ-NACE):

- CZ-ISCO vychází ze standardu ISCO-08² a je založena na druhu vykonávané práce (pracovní místo) a úrovni dovedností. Klasifikace rozděluje zaměstnání do hlavních tříd, které obsahují třídy, skupiny, podskupiny a kategorie.

²International Standard Classification of Occupations

- CZ-NACE byla vytvořena na základě klasifikace NACE³, kterou používá Evropská unie. Ekonomickými činnostmi se rozumí výroba určitého produktu nebo služby. NACE ovšem neurčuje přesné kategorie, ale poskytuje pravidla pro jejich tvorbu, na jejichž základě vznikla CZ-NACE. Velkým přínosem této klasifikace je možnost srovnávat naše data se všemi zeměmi EU.

Dále ČSÚ používá například Klasifikaci produkce (CZ-CPA), stavebních děl (CZ-CC), postavení v zaměstnání (CZ-ICSE), vzdělání (CZ-ISCED 2011), oborů vzdělání (CZ-ISCED-F 2013), územních statistických jednotek (CZ-NUTS) nebo zemí (CZ-GEONOM).

Číselníky poskytují seznamy unikátních kódů a jejich významů a patří mezi ně např. Číselník zemí (CZEM), Číselník měn a fondů (ČMF) nebo Číselník obcí (CISOB). Zároveň zde najdeme i číselníky, které poskytují kódy pro jednotlivé klasifikace.

Mezi ukazatele se řadí všechny statistikou sledované jevy a ve Statistickém metainformačním systému jsou náležitě popsány.

2.3 Kategorizace dat

ČSÚ v dnešní době sbírá a poskytuje data z mnoha oblastí. U většiny z nich zobrazuje základní graf nebo tabulku pro rychlý přehled a zároveň přikládá odkazy na přehledy konkrétních dat, metodiky, analýzy nebo související informace. V poskytnutých datech najdeme například časové řady, grafy, publikace nebo odkaz na vybrané tabulky z Veřejné databáze. Následující seznam popisuje oblasti dat, kterými se ČSÚ zabývá, a jejich významné ukazatele.

- Ceny, inflace: míra inflace, ceny výrobců, ceny zahraničního obchodu, ceny nemovitostí
- Cestovní ruch: kapacity a návštěvnosti v hromadných ubytovacích zařízeních
- Doprava, informační a komunikační činnosti] : indexy tržeb, přeprava věcí a osob
- Finanční hospodaření: příjmy a výdaje územních rozpočtů
- HDP, národní účty
- Informační technologie: produkce a využívání informačních technologií
- Konjunkturální průzkumy: saldo indikátoru důvěry
- Kriminalita, nehody: registrované a objasněné trestné činy, nehody, požáry
- Kultura: divadla, muzea, galerie, knihovny, televizní a rozhlasová vysílání
- Lesnictví: kategorizace a vlastnictví lesů, těžba dřeva, zalesňování
- Mzdy a náklady práce: průměrná mzda, podíly zaměstnanců, placený čas
- Obchod, pohostinství, ubytování: indexy tržeb
- Obyvatelstvo: počet narození, úmrtí, stěhování a věkové složení obyvatel
- Organizační statistika: přehled ekonomických subjektů

³Zkratka NACE je odvozena z francouzského názvu „Nomenclature générale des Activités économiques dans les Communautés Européennes“

- Příjmy, výdaje a životní podmínky domácností
- Průmysl, energetika: index a tržby z průmyslové činnosti
- Sčítání lidu, domů a bytů: výsledky sčítání
- Služby: indexy tržeb
- Sociální zabezpečení: vyplacené dávky, důchody, zařízení sociálních služeb
- Stavebnictví, byty: stavební práce, stavební zakázky, zahájené a dokončené byty
- Území, sídelní struktura: počty územních jednotek, zařazení obcí
- Věda, výzkum a inovace: finanční a lidské zdroje, výsledky výzkumu a vývoje
- Volby: volební účast, výsledky voleb
- Vzdělávání: počty studentů, učitelů, zaměstnanců a škol
- Zahraniční obchod: obraz, vývoz, dovoz a bilance zboží
- Zaměstnanost, nezaměstnanost: míra zaměstnanosti, uchazeči o zaměstnání
- Zdravotnictví, pracovní neschopnost: nemocnice, léčebny, lékárny, nemocensky pojištění nebo dny pracovní neschopnosti
- Zemědělství: počty zvířat, zemědělských plodin, výroba
- Životní prostředí: chráněná území, produkce odpadů, vodovody, kanalizace

Vedle zmíněných oblastí ČSÚ také nabízí tzv. *Průřezové statistiky*. Řadí do nich informace o cizincích, o seniorech, genderové statistiky, souhrnná data o České republice, makroekonomické údaje nebo mezinárodní data.

2.4 Výběr dat

Po prozkoumání všech oblastí, které ČSÚ poskytuje, jsem vybrala data z oblasti zaměstnanosti, nezaměstnanosti, průměrných mezd, vzdělání a kriminality. Ve Veřejné databázi jsou zobrazena jen v podobě tabulek a nejsou k dispozici podrobnější přehledy a možnosti srovnání. Přitom při vhodném zobrazení tato data mohou nabídnout přínosné informace například pro studenty středních škol, kteří se rozhodují, kterým oborem se budou v budoucnu zabývat. Informace o uplatnění a finanční situaci v daném oboru mohou jejich rozhodnutí výrazně ovlivnit.

2.4.1 Zaměstnanost, nezaměstnanost

Ke zpracování dat o zaměstnanosti a nezaměstnanosti se využívají především výsledky z Výběrového šetření pracovních sil (**VŠPS**), jehož cílem je pravidelné získávání informací o trhu práce. Tyto informace jsou získávány pomocí dotazníků v náhodně vybraném vzorku domácností. Obsah a forma dotazníku odpovídá standardům Evropské unie, proto jsou výsledky srovnatelné se všemi státy EU.

V dostupných tabulkách ve Veřejné databázi najdeme ukazatele, které jsou dále členěny podle atributů rok, čtvrtletí, kraj, pohlaví a některé i podle zaměstnání CZ-ISCO, postavení v zaměstnání, odvětví ekonomické činnosti CZ-NACE nebo vzdělání.

- **Počet zaměstnaných osob:** Počet všech osob patnáctiletých a starších, které v referenčním týdnu pracovaly aspoň 1 hodinu nebo nebyly v práci, ale měly formální vztah k zaměstnání. Do této kategorie patří všichni placení zaměstnanci, podnikatelé, profesionální příslušníci armády a osoby na mateřské dovolené, které před nástupem pracovaly.
- **Počet nezaměstnaných osob:** Všechny osoby patnáctileté a starší, které splňují tři podmínky. Nejsou zaměstnané, jsou připravené k nástupu do práce nejpozději do 14 dnů a hledají aktivně práci. Tento ukazatel je pouze odhad z **VŠPS**.
- **Pracovní síla v národním hospodářství:** Součet jen zaměstnaných osob v hlavním zaměstnání nebo naopak ekonomicky neaktivní osoby, které nepatřily do kategorie pracovních sil.
- **Obecná míra nezaměstnanosti:** Podíl počtu nezaměstnaných na celkové pracovní síle.
- **Míra ekonomické aktivity:** Podíl počtu zaměstnaných na počtu všech osob patnáctiletých a starších.
- **Míra zaměstnanosti:** Podíl počtu zaměstnaných ve věku 15–64 let na počtu všech osob v této věkové skupině.
- **Pracovní místa v evidenci úřadu práce.**
- **Uchazeči o zaměstnání v evidenci úřadu práce.**

2.4.2 Průměrné mzdy

V této kategorii jsou dva nejdůležitější ukazatele a to průměrná hrubá měsíční mzda a počet zaměstnanců. Jejich společnými atributy jsou rok, čtvrtletí, kraj, pohlaví, klasifikace zaměstnání CZ-ISCO a odvětví ekonomické činnosti CZ-NACE. Porovnáním průměrné mzdy daného období se stejným obdobím minulého roku se počítá pokles nebo růst průměrné nominální mzdy vyjádřený v procentech. Růst nebo pokles průměrné reálné mzdy naopak vyjadřuje, o kolik % se zvýšil nebo poklesl podíl indexu průměrné nominální mzdy a indexu spotřebitelských cen za shodné období.

- **Průměrná hrubá měsíční mzda:** Podíl mezd připadající na jednoho zaměstnance evidenčního počtu za měsíc. Do mezd se započítávají i příplatky, doplatky ke mzdě, odměny a jiné složky mezd nebo platů.
- **Evidenční počet zaměstnanců:** Počet osob v pracovním, služebním nebo členském poměru k zaměstnavateli.
- **Medián mezd:** Hodnota mzdy nebo platu uprostřed mzdového rozdělení.

2.4.3 Vzdělávání

Tabulky pro tuto oblast jsou rozděleny podle druhu vzdělávání. Společnými atributy všech druhů jsou počet škol, tříd, žáků nebo studentů a učitelů. Dále nechybí rozlišení pohlaví studentů a učitelů, časové srovnání a rozdělení podle krajů. Specifické atributy pro jednotlivé druhy jsou vypsány níže:

- **Předškolní vzdělávání**
- **Základní vzdělávání**
- **Střední vzdělávání:** nově přijatí studenti a absolventi; obory gymnázií, odborných škol a nástavbového studia; forma studia; kvalifikace učitelů
- **Vyšší odborné vzdělávání:** nově přijatí studenti a absolventi; forma studia
- **Vysokoškolské vzdělávání:** poprvé zapsaní studenti a absolventi; školy veřejné a soukromé; forma studia; typ studijního programu; typ akademických pracovníků

2.4.4 Kriminalita, nehody

Oblast kriminality by mohla souviset s nezaměstnaností v dané oblasti a zároveň by mohla uživatele zajímat, pokud by například uvažoval o stěhování za novou školou nebo zaměstnáním.

V této oblasti ČSÚ v první řadě zobrazuje počty registrovaných a objasněných trestných činů a jejich roční srovnání. Ve veřejné databázi je k dispozici tabulka s pravomocnými rozhodnutími soudu a průměrnými délkami soudního řízení, dále tabulka s počty odsouzených osob za vybrané trestné činy a tabulka trestných činů a jejich objasněnost. Důležitými ukazateli jsou počty trestných činů, nehod a požárů v České republice, které jsou zobrazovány v tabulkách podle územního nebo časového srovnání. Jejich další atributy jsou v seznamu:

- **Trestné činy:** obecná kriminalita, hospodářská kriminalita, loupeže, vloupání do bytů a domů, znásilnění a vraždy
- **Nehody v silniční dopravě:** nehody pod vlivem alkoholu, usmrcené osoby, zraněné osoby, hmotná škoda
- **Požáry:** přímá škoda, uchráněné hodnoty, usmrcené osoby, zraněné osoby

2.5 Další zdroje dat

Statistická a jiná veřejná data samozřejmě poskytují i jiné organizace a to například Český úřad zeměměřický a katastrální, Český telekomunikační úřad, Česká správa sociálního zabezpečení, ministerstva ČR nebo města. Informace, které poskytují, jsou k dispozici na jejich stránkách.

Další data z oblasti zaměstnanosti jsou dostupná na stránkách Ministerstva práce a sociálních věcí (**MPSV**). Poskytuje zde statistická data v podobě tabulek, přičemž některé informace jsou k zobrazení přímo na webových stránkách a jiné jsou dostupné v souborech ke stažení. Oblasti dat, která MPSV poskytuje, jsou:

- Nezaměstnanost
- Absolventi škol a mladiství
- Rekvalifikace
- Nabídka a poptávka na trhu práce
- Uchazeči z EU, EHP a třetích zemí
- Zaměstnávání cizích státních příslušníků
- Statistika výdělků
- Statistická ročenka
- Výdaje na státní politiku zaměstnanosti

Podstatnou změnou v poskytování dat veřejné správy bylo vytvoření tzv. *Otevřených dat*⁴, která vychází ze standardů a technologií vytvořených na základě zkušeností prvních otevřených zemí (USA, UK) [3]. Otevřená data přes Národní katalog otevřených dat⁵ poskytují veřejnosti určité datové soubory online ve strojově čitelném a otevřeném formátu, které jsou opatřeny dokumentací a jsou dostupné ke stažení bez technických překážek.

Výhodou otevřených dat je především budoucí použití. Díky jejich surové podobě a snadné dostupnosti je pro vývojáře jednodušší vytvářet webové a mobilní aplikace, které využívají data z veřejné správy. Od roku 2013 pořádá Nadace OSF⁶ každoročně soutěž o nejlepší aplikace postavené nad otevřenými daty s názvem Společně otevíráme data. O rok později vznikl projekt Hackathon, kde účastníci vytvářejí aplikace dle zadání během 48 hodin.

Nicméně datových sad z oblastí, které jsem pro vizualizaci vybrala, v Otevřených datech moc nenajdeme. Obsahují jen datovou sadu o uchazečích o zaměstnání, podílu nezaměstnaných osob a počty obyvatel.

⁴<http://www.otevrenadata.cz>

⁵<http://www.data.gov.cz>

⁶<http://www.osf.cz>

Kapitola 3

Dashboard a vizualizace dat

Než se zaměříme na dashboard a jeho tvorbu, je dobré si uvědomit, proč chceme data vizualizovat. Proč je grafická reprezentace účinnější než textové zobrazení? Grafická reprezentace pomáhá zobrazovat více hodnot, přiměje uživatele vnímat spíše kontext dat a pomáhá interpretovat data tak, jak bylo zamýšleno. Také umožňuje lépe porovnávat hodnoty. [18]

Dashboard je nástroj, který klade důraz na vhodnou grafickou reprezentaci dat. Na internetu najdeme mnoho produktů pod názvem dashboard. Po jejich prostudování zjistíme, že dané dashboardsy nemají vždy stejné vlastnosti, ale shodují se v zobrazování dat v podobě grafů, tabulek nebo oblíbených měřidel jako na palubní desce auta.

Tato kapitola se zabývá definicí dashboardu, jeho klasifikací a následně charakterizuje jeho tvorbu.

3.1 Definice dashboardu

Vhodnou definici o tom, jak by měl dobře navržený dashboard vypadat, vytvořil Stephen Few [5]:

„A dashboard is a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance.“

Podle jeho definice musí správný dashboard splňovat několik pravidel. Zaprvé klade důraz na vizuální zobrazení dat. Dashboard by měl především obsahovat grafické prvky, protože pokud nehledáme konkrétní hodnoty, grafická reprezentace pomáhá pochopit kontext dat. K vytvoření takto efektivního zobrazení je však potřeba rozumět jak vizuální vnímání funguje.

Dalším bodem, který musí dashboard splňovat je zobrazování všech dat tak, aby se vešly na jednu obrazovku. Všechno potřebné by mělo být viditelné na první pohled bez nutnosti scrollování či klikání na nějaké podstránky pro nalezení důležité informace. Pro uživatele tak není těžké si všechny informace spojit a logicky uspořádat v hlavě.

Dále by měl dashboard zobrazovat nejnovější dostupná data. Pro některá podnikání je potřeba monitorovat data každou sekundu a pro jiná stačí pouze data za minulý měsíc. Důležité je, aby nám dashboard při neočekávané změně či výkyvu tento jev určitým způsobem ukázal. Není potřeba jej detailně popisovat, ale měl by uživatele upozornit na případný problém. Je proto důležité zmínit, že dashboard zobrazuje agregovaná či sumarizovaná data.

Dashboardsy najdeme často v oblasti podnikání. Slouží zde hlavně pro sloučení všech důležitých informací jako například přehledu o toku finančních prostředků, monitoringu

prodejů nebo stavu zásob. Většinou potřebná data pocházejí z různých zdrojů, proto je dashboard sjednocuje a zobrazuje všechna v rámci jednoho přehledu. Takový dashboard je nazván *Performance dashboard* a není již pouhým vizualizačním nástrojem. Podle Wayne W. Eckersona [4] je důležité jej vnímat jako plnohodnotný informační systém, který slouží pro lepší organizaci a řízení daného podniku. Definuje jej takto:

„A performance dashboard is a multilayered application built on a business intelligence and data integration infrastructure that enables organizations to measure, monitor, and manage business performance more effectively.“

Performance dashboard obsahuje aplikace pro monitorování, analýzu a řízení. Navíc poskytuje tři pohledy pro zobrazení informací a to sumarizovaný, multidimenzionální a detailní pohled. Jelikož se již nejedná o pouhou vizualizaci dat, je někdy nazýván performance management systém, aby nebyl mylně spojován s dashboardem.

3.2 Klasifikace dashboardů

Přestože máme definici pro dashboardy, stále budou hodně odlišné v datech, která zobrazují, a způsobu používání. V tabulce 3.1 najdeme seznam proměnných, podle kterých lze dashboardy klasifikovat. Dělení dashboardů na strategické, analytické a operační je jedno z nejpoužívanějších.

Proměnné	Hodnoty
Role	Strategické, analytické, operační
Typy dat	Kvantitativní, kvalitativní
Datové domény	Prodej, finance, marketing, výroba, lidské zdroje
Typy měření	Balanced Scorecard (např. KPI), Six sigma
Rozpětí dat	Celý podnik, oddělení, individuální
Aktualizace dat	Měsíční, týdenní, denní, hodinové, real-time
Interaktivita	Statické zobrazení, interaktivní zobrazení (filtry)
Způsoby zobrazení	Převážně grafické, převážně textové, obojí
Portálové funkce	Funkce získání dodatečných dat, bez funkce

Tabulka 3.1: Možnosti klasifikace dashboardů podle [5]

Strategická role

Strategické dashboardy se používají pro rozhodování o firmě, zahrnují přehledy o současné situaci, ale také umí zobrazovat předpovědi. Najdeme zde i další ukazatele, porovnávání dat nebo krátkou historii. Problém nastává, když je dat moc a uživatel v nich nenachází ta důležitá. Vzhled těchto rozhraní by měl být jednoduchý a soustředit se na přehlednost dat. Pro strategii nejsou potřeba real-time data ani interaktivní zobrazení.

Analytická role

Dashboardy, které slouží k analytickému účelu, musí být komplexnější než ty strategické. Není zde sice potřeba sledovat real-time data, ale už nestačí jen neinteraktivní přehledy. Analytické dashboardy by měly obsahovat například filtry, grafy s časovým vývojem, možnosti jiného zobrazení dat a odkazy na zjištění detailů o dané problematice, protože analytik

potřebuje ke zjišťování informací širší kontext než mu samotný dashboard může na první pohled ukázat.

Některé analytické dashboardsy využívané ve firmách poskytují manažerům i možnost předpovědí (například co se stane, pokud zvýšíme ceny o x procent). Také umožňují vytváření vlastních pohledů a výpočtů na datech. Uživatelé tak nemusí spoléhat na výstupy dat, které připravil někdo jiný, ale mohou si vytvářet vlastní. Tato funkce většinou souvisí se sdílením vlastních výstupů například mezi ostatními pracovníky [4].

Do této kategorie se řadí i dashboard pro vizualizaci statistických dat, který je vytvářen v rámci této práce.

Operační role

Do poslední kategorie patří operační dashboardsy, které slouží k monitorování aktuálních dat. Dashboard musí dbát na dynamičnost a vysokou frekvenci aktualizace dat. Vyznačuje se především jednoduchým zobrazením dat a hlavně zvěrazňováním kritických hodnot. Na ně musí uživatel co nejrychleji reagovat a většinou se u nich zobrazují i všechny podrobnosti o problému.

3.3 Tvorba dashboardu

Před vytvářením dashboardu je důležité nejdříve rozhodnout, zda je právě dashboard nejvhodnějším nástrojem nebo naopak je lepší zvolit jiné zobrazení (např. balanced scorecard nebo modifikace stávajícího řešení). K rozhodnutí pomůže stanovení účelu, díky kterému následně můžeme zjistit, jaká data budeme zobrazovat [7].

Tři základní fáze procesu tvorby dashboardu jsou odpověďmi na následující otázky [12]:

1. Jaké informace?
2. Pro koho?
3. Jak je prezentovat?

Výběr dat je důležité kritérium pro vytvoření dashboardu, protože by měl zobrazovat právě data, která jsou potřebná. S tím souvisí stanovení vhodných metrik, které je podmíněno identifikací budoucích uživatelů, jejich požadavků a informací, které jsou pro ně důležité. Následně se navrhuje způsob zobrazení, který klade důraz na design a rozvržení prvků, výběr grafů, typografii nebo výběr správných barev.

3.3.1 Identifikace uživatelů

Identifikace budoucích uživatelů a jejich požadavků vyžaduje získání dat pomocí kvantitativního nebo kvalitativního výzkumu. Existuje mnoho metod, které se průzkumem uživatelů zabývají. Metody, které byly následně využity, jsou popsány níže.

Dotazník

Jednou z nejpoužívanějších metod sběru dat je dotazník, který řadíme do kvantitativního výzkumu. Dotazník slouží pro ověření hypotéz a obsahuje uzavřené otázky, ale také otevřené, které se již řadí do kvalitativního výzkumu. Jeho výhodou je rychlý sběr dat na velkém počtu respondentů [6].

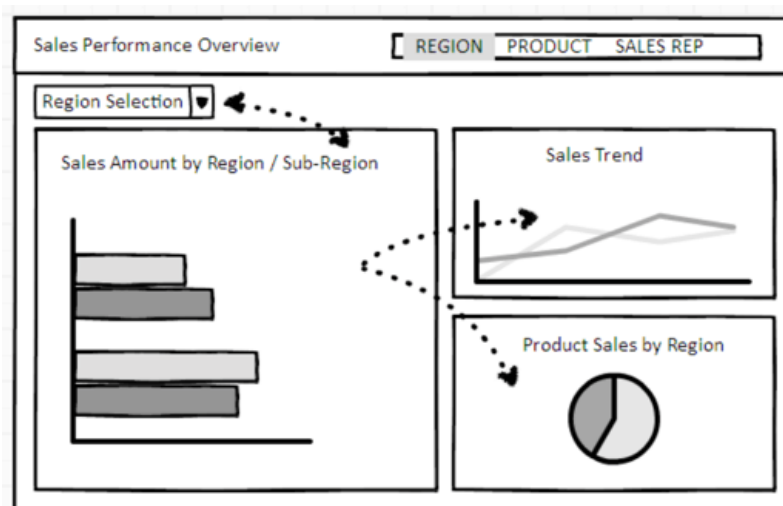
Persony

Vytváření person se řadí do kvalitativního výzkumu. Persona je fiktivní osoba, která reprezentuje typické uživatele systému a pomáhá lépe pochopit potřeby, zkušenosti a cíle uživatele. Díky nim je možné vžít se do uživatelské role a dosáhnout lepšího výsledku.

Prototypování

Prototypování můžeme rozdělit na několik metod. První z nich je skicování, kdy jsou jednotlivé prvky systému nakresleny a ručně popsány. Díky tomu je možné si částečně představit budoucí systém a následně probírat první nápady s uživateli.

Další metodou je vytvoření tzv. *drátěného modelu* (anglicky *wireframe*), který slouží k vizualizaci návrhu rozložení prvků na stránce. Většinou je již v digitální podobě, ale obsahuje jen jednoduché obrazce a text. Příklad drátěného modelu zaměřeného na dashboard můžete vidět na obrázku 3.1.



Obrázek 3.1: Příklad drátěného modelu zaměřeného na dashboard¹

Předchozí dvě metody vedou k vytvoření tzv. *prototypu*, který již obsahuje interaktivní prvky a často se velmi podobá finální verzi. Díky prototypu je možné zjistit další specifické požadavky uživatelů a odhalit případné chyby ještě před implementací systému.

3.3.2 Stanovení metrik

Zvolení správných metrik pomáhá soustředit uživatelovu pozornost na kritické oblasti. Metriky jsou důležitým faktorem a významně pomáhají například vedení firmy ke zlepšení strategie a posunutí novým směrem. Pro shromáždění požadavků se využívají rozhovory nebo dotazníky s výkonnými pracovníky o strategii a cílech firmy [4].

Vhodná metrika by měla vyvolávat určitou reakci při změně hodnoty, být běžně interpretována a její vygenerování by mělo být všem jasné. Měli bychom se vyvarovat metrikám, které kombinují mnoho faktorů. Také zdroj dat by měl být lehce dostupný a věrohodný [8].

¹Převzato z <https://visualbi.com/blogs/business-intelligence/dashboards/importance-of-wireframing/>

Metriky, které měří, jak dobře organizace vykonává operační, taktické nebo strategické aktivity, se nazývají *klíčové identifikátory výkonu* (Key Performance Indicators - KPI). Můžeme je rozdělit na *leading indicators* (řídící) a *lagging indicators* (vyplývající z minulosti). Stanovení lagging indicators je často jednodušší (např. spokojenost zákazníků), ale leading indicators jsou klíčovou hnací silou, která předpovídá budoucí výsledky (např. vyřešení problému během první návštěvy) [4].

3.3.3 Vizualizace

Pokud už víme, co budoucí uživatelé systému nejvíce potřebují a máme zvolené metriky, můžeme se věnovat designu a struktuře dashboardu. Existují určitá pravidla pro tvorbu uživatelských rozhraní (**UI** z anglického user interface), která je potřeba dodržovat. Tato pravidla vychází z principů vizuálního vnímání. Při jejich zkoumání se můžeme zaměřit na ukládání informací do paměti a jejich rychlé zpracování. [5].

Grafické prostředky

Výběr grafů, měřidel, tabulek či jiných prostředků, které budou vybrané informace vhodně vizualizovat, je důležitou a často složitou činností. Stephen Few se ptá:

„Měla by být informace zobrazena v podobě textu, graficky nebo pomocí obou možností?“

Každá možnost má své výhody i nevýhody. Zobrazování individuálních hodnot nepotřebuje žádnou grafickou reprezentaci a pro hledání konkrétní hodnoty ve větším množství dat je stále lepší využít tabulku než graf. Nicméně pro srovnávání více informací stejného typu je lepší využít graf. Dashboardy proto využívají obě možnosti zobrazování [5].

V některých případech je jasně dané, jaký grafický ukazatel se použije, ale jindy to vyžaduje určitý stupeň kreativity a vynalézavosti [12]. Vzhledem k tomu, že hlavní princip dashboardu je zobrazení všech informací na jedné obrazovce, je mnohdy potřeba všechny informace dostat do malého prostoru a zároveň je zobrazit efektivně.

Hlavním grafickým prostředkem pro vizualizaci dat jsou grafy. Mezi nejznámější typy grafů, které se u dashboardů používají, patří:

- Sloupcové grafy
- Skládané grafy
- Spojnicové grafy
- Bodové grafy
- Bullet grafy
- Sparkline
- Box plot
- Treemap

Seznam sestavil Stephen Few [5] a nezařadil do něj koláčové grafy, protože v tomto typu grafu jsou hodnoty hůře porovnatelné.

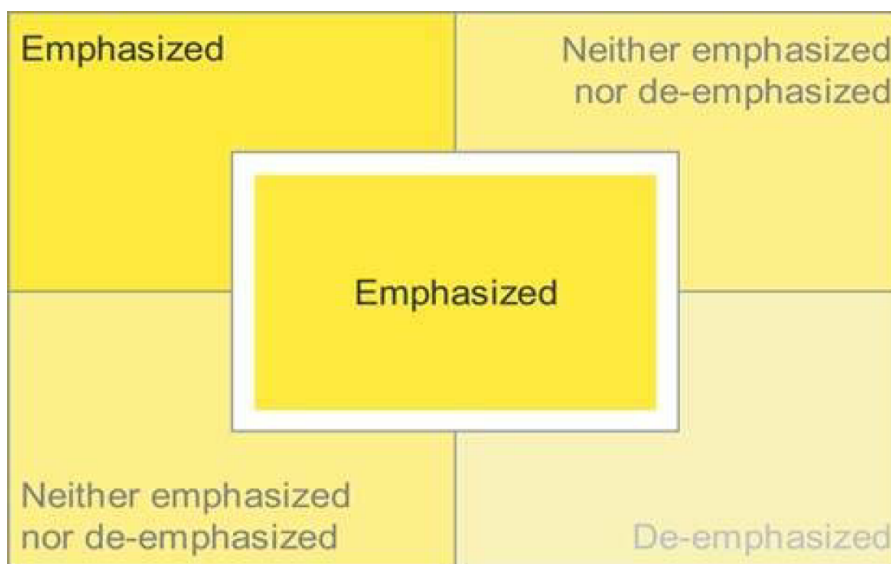
Dalším využívaným prostředkem jsou ikony, které slouží k upozorňování nebo k vizualizaci určitých změn. Můžeme ještě zmínit obrázky, tabulky, mapy nebo různé kombinace grafů.

Na závěr je dobré si uvědomit, že vybrání například grafů s 3D efekty či různými barevnými a jinými dekoracemi se do dashboardu opravdu nehodí. Důležité je, co grafy zobrazují, jak spolu souvisí, jsou jednoduché a přehledné [4].

Struktura

Běžnou praxí bylo zobrazení více grafů pomocí mřížky, kdy nebylo zřejmé, jestli spolu grafy nějakým způsobem souvisí či nikoliv. Vybrané grafy a jiné vizualizační prostředky by ale měly být zobrazovány podle jejich souvislostí tak, aby uživatel rychle pochopil, co grafy zobrazují. Grafy s podobným významem je vhodné například slučovat do skupin [8].

Vzhledem k nutnosti zobrazit všechna data na jedné obrazovce je důležité vědět, jak uživatel celou obrazovku vnímá a ve které části hledá nejdůležitější informace. Takové informace je nejlepší zobrazit v levém horním rohu nebo na střed (viz obrázek 3.2).



Obrázek 3.2: Oblasti vizuálního vnímání obrazovky [5]

Barevnost

Využití barev má své silné i slabé stránky. Naše oko je přizpůsobeno vnímání kontrastu barev a schopnost odlišit barvy od sebe závisí na jejich prezentaci. Také je potřeba zmínit, že ne každý uživatel má k dispozici monitor s širokým barevným spektrem a dříve či později může daný systém využívat i barvoslepý člověk [10].

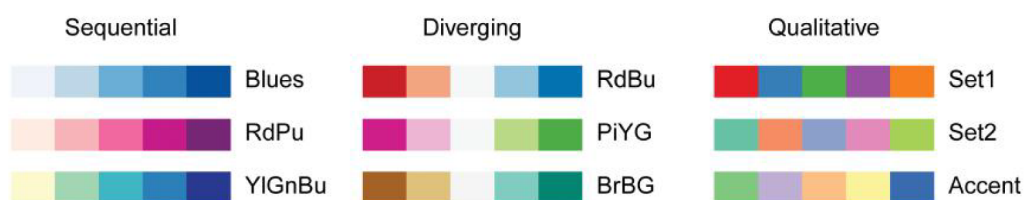
Pravděpodobně nejdůležitějším aspektem při výběru vhodných barev je zjištění, že barvy jsou vnímány relativně vzhledem k okolním barvám. Například na obrázku 3.3 jsou zobrazeny obdélníky stejné barvy, ale na různobarevném pozadí se stejné nezdají [9].

Barvy je dobré rozlišit podle sytosti, jasů a také odstínu, aby byly od sebe dobře rozzeatelné. Pokud se převedou do odstínů šedi, jejich rozdíl by měl být stále patrný [10]. Na internetu již existuje mnoho nástrojů pro správný výběr kombinací barev, například



Obrázek 3.3: Iluze vnímání odstínů barev [9]

ColorBrewer². Barevné palety můžeme rozdělit podle zobrazovaných dat do tří typů na sekvenční, divergentní a kvalitativní (viz obrázek 3.4).



Obrázek 3.4: Rozdělení barevných palet podle typu zobrazovaných dat (ColorBrewer)

3.3.4 Testování

Nedílnou součástí procesu tvorby nového systému je testování použitelnosti. Je doporučeno provádět testování již při vývoji a brzy odhalit chyby. Test použitelnosti se provádí na jednotlivcích, kterým je prezentován nový systém nebo jeho prototyp, a mají za úkol buď identifikovat systém jako takový, pokud ho ještě neznají, nebo splnit několik úkolů. Steve Krug doporučuje provádět více testů s malým počtem uživatelů než jeden s více uživateli. Po jednotlivých testech je možné opravit problémy a při dalším testu tak narazit na další, které dosud nemohly být objeveny [11].

²Dostupné na <http://colorbrewer2.org>

Kapitola 4

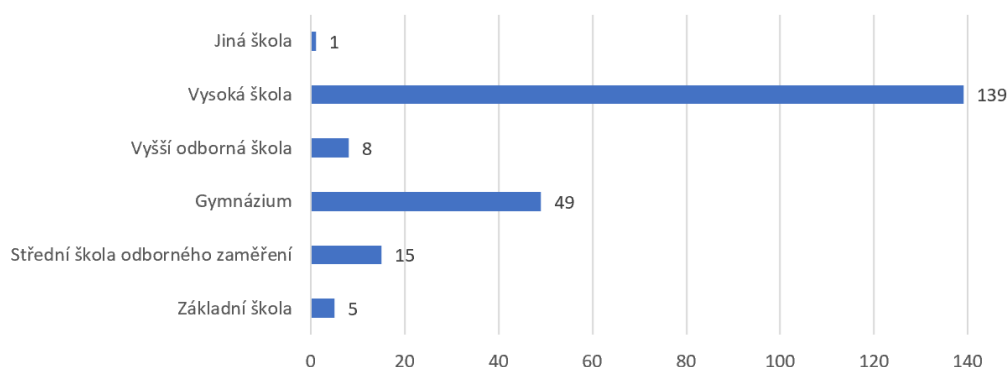
Analýza uživatelů a požadavků

V této kapitole se budeme zabývat průzkumem uživatelů pro zjištění jejich požadavků a potřeb na systém. Zjistíme, jaké informace je zajímají a co by jim dashboard vizualizující tyto informace mohl přinést. Také budeme analyzovat současné možnosti zobrazení potřebných informací. Je tedy potřeba zodpovědět následující otázky:

1. Kdo je naše cílová skupina?
2. Jaké informace je zajímají?
3. Kde takové informace najdou dnes?
4. Můžeme je zobrazit lépe?

4.1 Průzkum uživatelů

Pro získání informací o uživateli byl vytvořen krátký dotazník. Celkem jej zodpovědělo 271 respondentů v různých věkových kategoriích, kdy 60 % z nich mělo 19-25 let a 80 % byli studenti, na které bude systém hlavně cílit. Poměr odpovědí studentů na otázku: „*Jakou úroveň vzdělání aktuálně studujete?*“, je zobrazeno v grafu 4.1.



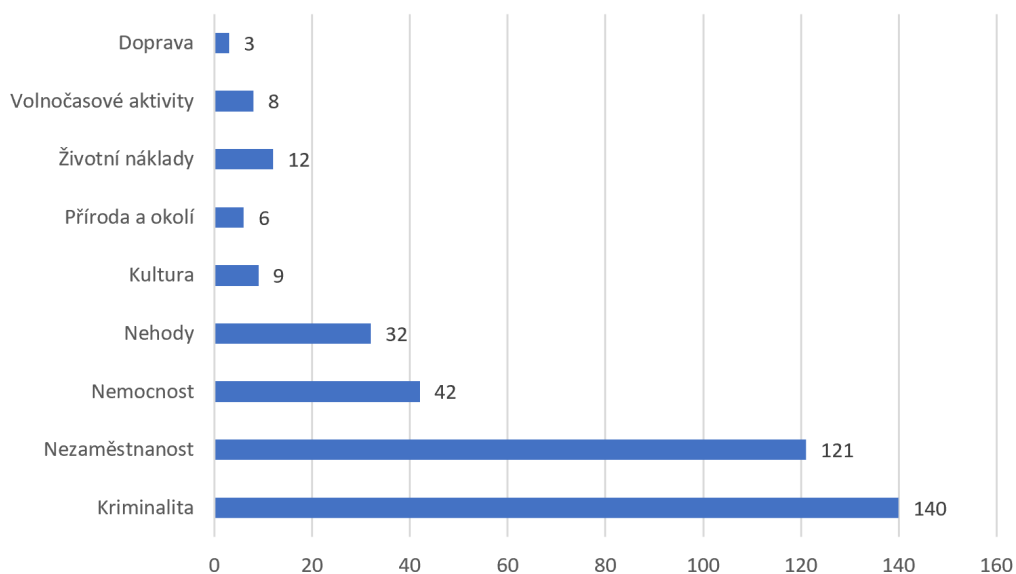
Obrázek 4.1: Rozdělení studentů podle úrovně vzdělání v rámci dotazníku.

Z dotazníku bylo dále zjištěno, že 71 % studentů je již rozhodnutých, jaké povolání by chtěli vykonávat. Nejdůležitějším kritériem při výběru povolání je pro ně zájem o daný obor

(90 %) a druhým kritériem je platové ohodnocení (76 %). Některé respondenty by ale také ovlivnily předchozí zkušenosti v oboru, vzdálenost práce od bydliště a uplatnění v oboru.

Další otázka z dotazníku směřovala na téma informovanosti o tom, jak je dané povolání žádané. Nadpoloviční většina (58 %) respondentů se již předem o žádanosti povolání informovala. Ze zbývajících dotázaných 27 % projevilo o tyto údaje alespoň zájem. Ze skupiny pracujících, což bylo 20 % respondentů, odpověděla pětina z nich na otázku, zda uvažují o změně zaměstnání, kladně.

Ukázalo se, že 68 % respondentů by při rozhodování o zaměstnání ovlivnily údaje o mzdách zobrazené podle lokality. Za lepší práci by se přestěhovalo 13 %, nicméně 67 % možné přestěhování nezavrhovalo. Na otázku: „*Jaké informace by Vás ještě při výběru povolání zajímaly o dané lokalitě?*“, mohli respondenti vybírat z předem připravených odpovědí a také přidat vlastní. Respondenti by uvítali informace o kriminalitě, nezaměstnanosti, nehodách nebo např. cen nemovitostí v dané lokalitě. Všechny čtenější odpovědi jsou zobrazeny v grafu 4.2.



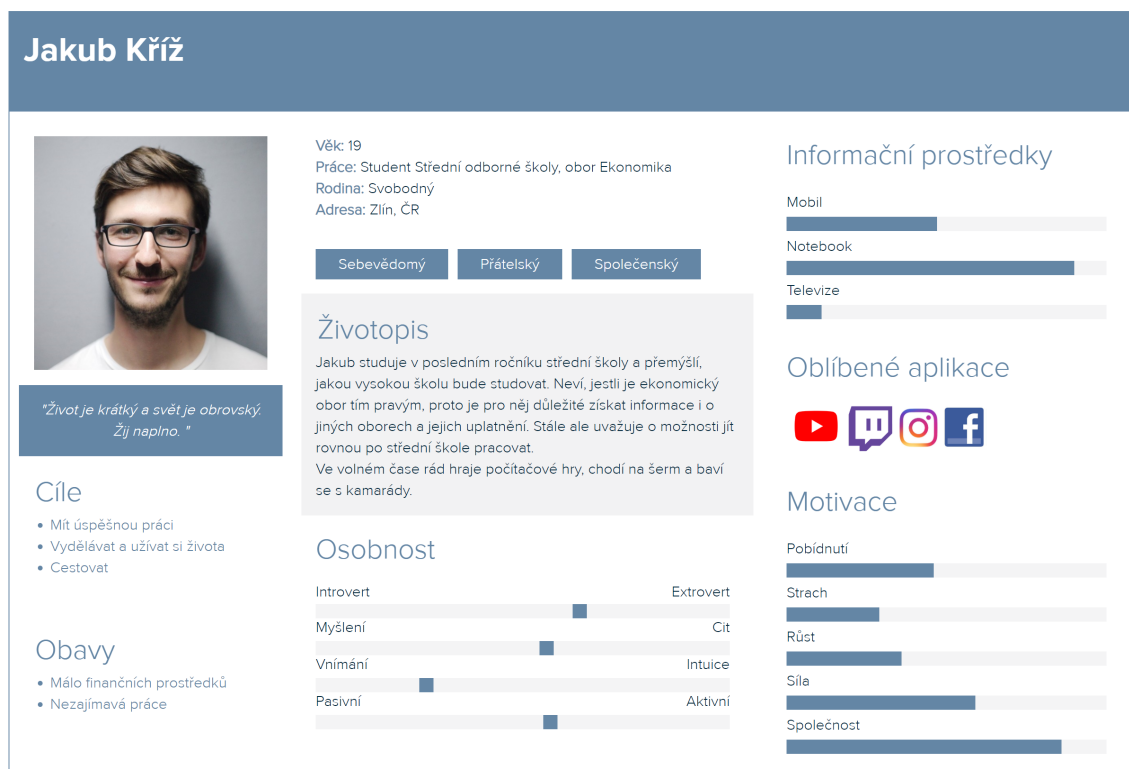
Obrázek 4.2: Odpovědi respondentů na otázku, jaké informace o dané lokalitě by je zajímaly.

Na konci dotazníku měli respondenti možnost uvést webové stránky či jiné zdroje, díky kterým získávají data o zaměstnanosti a průměrných platech v České republice. Nicméně 93 % z nich uvedlo, že se s takovými stránkami nesešli.

Z dotazníku vyplývá, že respondenty informace o průměrných platech, nezaměstnanosti nebo uplatnění v oboru zajímají, ale dosud je moc nevyhledávali. Maximálně si zjišťovali průměrné platy v daném oboru například na webových stránkách pro hledání práce.

Pro upřesnění představy o typických uživateli bylo navrženo několik person. Příklad jedné z nich je na obrázku 4.3. Typickým uživatelem informačního systému bude student střední školy, který přemýšlí, co bude studovat na vysoké škole nebo jaké zaměstnání si zvolí. Pokud si dané studium nebo povolání vybere, měl by mít možnost získání informací o uplatnění v oboru. To zahrnuje například informace o průměrném platu, o počtech uchazečů o toto zaměstnání a počtech nabízených míst.

Dalšími uživateli systému mohou být i studeni základní školy, kteří se tak mohou seznámit se situací v zaměstnanosti v České republice, nebo naopak studenti vysokých škol,



Obrázek 4.3: Příklad jedné z vytvořených person.

kterí si už obor vybrali a mohou jej srovnávat s ostatními. V neposlední řadě je systém využitelný i pro zaměstnané osoby ať už v porovnání jejich oboru s ostatními nebo zjištění poptávky a nabídky po tomto oboru. Pokud uživatel například hledá práci a řeší, zda by se měl kvůli ní přestěhovat, systém by mu mohl pomoci s porovnáním informací o krajích a okresech v České republice.

4.2 Současná dostupnost dat

Informace, které uživatele zajímají, jsou dnes na internetu jednoduše dostupné. Problémem může být jejich vyhledávání nebo právě vizualizace. Jednotlivé ukazatele (ne)zaměstnanosti, průměrných platů nebo uplatnění absolventů jsou často probírány ve zpravodajstvích nebo se jim věnují různé webové portály a některé z nich jsou níže uvedeny. Data jsou často zobrazena pomocí tabulek nebo zbytečně složitě interpretována.

Data, která poskytuje Statistický úřad, jsem rozepsala ve 2. kapitole. Lidé zde mohou najít mnoho informací, které by je mohly zajímat, ale problémem by mohlo být jejich zobrazení. Při prvním pohledu na webovou stránku úřadu¹ jsou zobrazeny hodnoty počtu obyvatel, průměrné mzdy, míry inflace, hrubém domácím produktu, průmyslové produkci a stavební produkci. Dále je možnost výběru kraje, po kterém se zobrazí podobný žebříček hodnot v rámci daného kraje. Pokud chce uživatel například získat více informací o průměrné mzdě, je mu zobrazen graf s jejím vývojem v čase. Další informace jako průměrné

¹<https://www.czso.cz/>

mzdy podle krajů nebo podle zaměstnání jsou dostupné až při dalším hledání. Podrobné informace jsou k zobrazení v tabulkách ve Veřejné databázi.

Průměrným mzdám se věnuje například portál [Platy.cz](https://www.platy.cz/)², který za informaci o platu uživatele poskytuje další podrobná data. Podrobné statistiky výdělků jsou také dostupné na portálu Ministerstva práce a sociálních věcí, které nabízí stažení rozsáhlých publikací pro jednotlivé kraje a období³.

Vizualizace míry nezaměstnanosti v čase je například na webové stránce [Peníze.cz](https://www.penize.cz/)⁴, kde jsou nabídnuty i jiné zajímavé grafické přehledy jako vývoj cen bytů, cen benzínu nebo průměrných mezd. Statistikám o uplatnění v oboru se věnuje portál [VysokeSkoly.com](http://www.vysokeskoly.com/uplatneni-absolventu/)⁵, který zobrazuje seznam všech vysokých škol a jejich fakult, u kterých uvádí procenta nezaměstnaných absolventů.

²<https://www.platy.cz/>

³<https://portal.mpsv.cz/sz/stat/vydelky>

⁴<https://www.penize.cz/nezamestnanost>

⁵<http://www.vysokeskoly.com/uplatneni-absolventu/>

Kapitola 5

Návrh systému a získání dat

Po zjištění, která data jsou k dispozici a které informace cílovou skupinu zajímají, bylo nutné stanovit metriky a grafické prvky dashboardu. Návrh se řídil podle principů tvorby dashboardu definovaných v kapitole 3.

5.1 Stanovení metrik

Většina dat, která byla pro vizualizaci vybrána, je dostupná nejen pro Českou republiku, ale i pro kraj, okres nebo pro dané zaměstnání. Aby měl uživatel možnost zobrazení dat např. pro konkrétní kraj, bylo nutné navrhnout filtrování. Vybrané ukazatele a jejich metriky jsou znázorněny v tabulce 5.1:

Ukazatel	ČR	Kraj	Okres	Zaměstnání
Počet obyvatel	✓	✓	✓	
Počet ekonomicky aktivních a neaktivních obyvatel	✓	✓		
Průměrná mzda	✓	✓		✓
Počet a rozdělení zaměstnaných osob podle pohlaví	✓	✓		✓
Počet a rozdělení nezaměstnaných osob podle dosaženého vzdělání	✓	✓		
Počet uchazečů o zaměstnání a volných míst	✓	✓	✓	✓
Počet studentů vysokých škol, absolventů a poprvé zapsaných	✓			
Počet nezaměstnaných absolventů vysokých škol	✓	✓	✓	
Počet trestných činů, nehod, požárů	✓	✓	✓	

Tabulka 5.1: Vybrané ukazatele a jejich metriky

Jak bylo zmíněno v kapitole 2, některá vybraná data ze Statistického úřadu jsou členěna podle klasifikace zaměstnání CZ-ISCO i klasifikace ekonomických činností CZ-NACE. Do dashboardu byla zvolena jen klasifikace zaměstnání, protože je podrobnější a je využívána i na Ministerstvu práce a sociálních věcí, ze kterého bude dashboard také čerpat data.

Vzhledem k tomu, že ne všechna data jsou dostupná pro kraje, okresy i zaměstnání, dashboard by se rozdělil do několika částí, které by se měnily v závislosti na vybraném filtru.

Některé prvky by zůstaly zobrazeny jen pro Českou republiku, jiné by se zobrazily pro daný kraj, další zase pro dané zaměstnání, ale mohli bychom najít i ukazatele, které obsahují data pro obě možnosti. V tom případě by bylo jednodušší umožnit filtrování pro každý grafický prvek zvlášť v závislosti na jeho možnostech, ale to by mohlo být komplikované a nepřehledné.

Druhým a následně zvoleným řešením bylo navržení několika podobných obrazovek pro každý typ filtru zvlášť. Každá obrazovka by měla reprezentovat jeden konkrétní pohled na data, aby bylo možné se lépe zaměřit na rozmístění jednotlivých prvků.

5.2 Návrh obrazovek

Proces návrhu obrazovek se neobešel bez několika návrhů na papíře, kde bylo načrtnuto rozvržení jednotlivých prvků hlavní obrazovky a následné tvorby drátěných modelů (anglicky *wireframe*). Hned zpočátku bylo rozhodnuto, že vpravo nahoře bude webová stránka obsahovat dva filtry, podle kterých si uživatel bude moct zobrazit informace jen pro daný kraj nebo dané zaměstnání. Filtrování okresu nebylo nutné, protože všechny informace o okresech by měly být zobrazeny již v obrazovce věnované kraji. Byly tedy navrženy tři obrazovky.

Dashboard pro ČR

Dashboard na hlavní obrazovce by měl zobrazovat data pro Českou republiku a nabízet určitá srovnání krajů a zaměstnání. Důležitým ukazatelem je průměrná mzda, proto by měla být zobrazena v levém horním rohu. Uživatel může průměrné mzdy lépe vnímat, pokud budou zobrazeny na mapě a rozlišeny barevně. Počet obyvatel, ekonomická aktivita, počet zaměstnaných a nezaměstnaných budou také v první části dashboardu zobrazeny buď číselně nebo pomocí koláčových grafů. I když před tímto typem grafu Stephen Few [5] varuje, myslím si, že pro vizualizaci dvou hodnot (jako např. pohlaví zaměstnaných) budou vhodné.

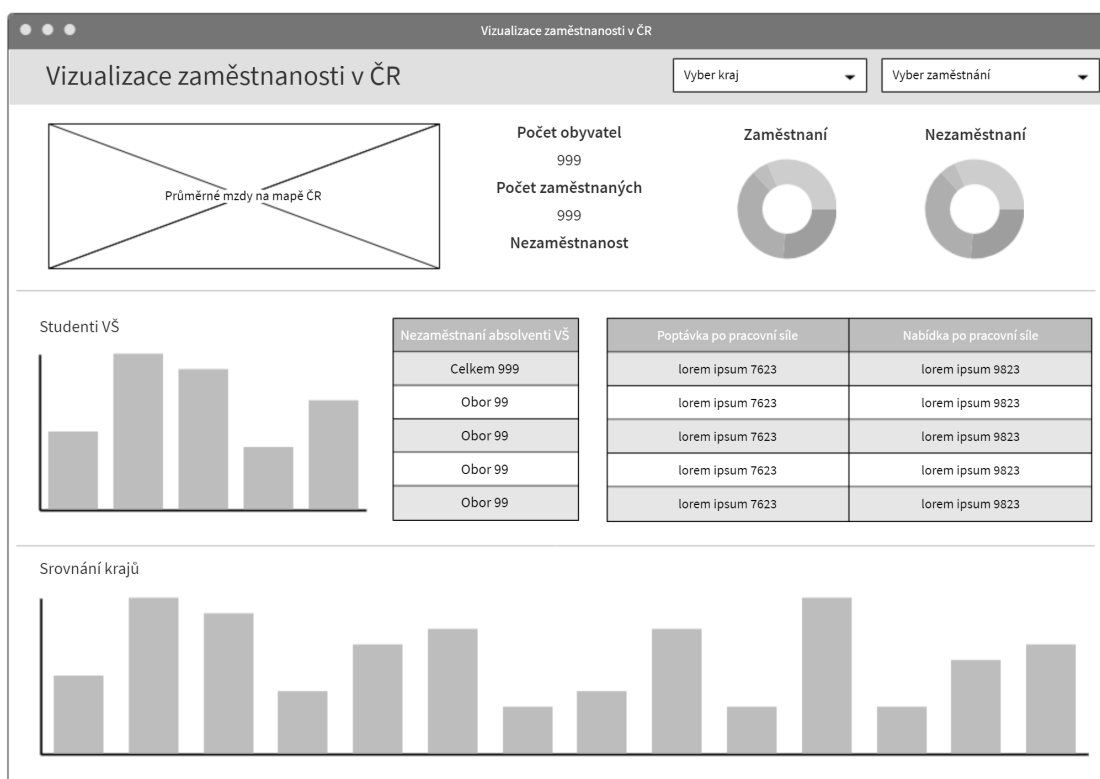
Druhá část bude zaměřena na studenty vysokých škol, nezaměstnanost absolventů a nabídne srovnání uchazečů o zaměstnání a volných míst podle zaměstnání. Studenti se mohou rozlišovat na poprvé zapsané a absolventy, proto byl navržen skládaný graf a vývoj v čase. Druhů zaměstnání je mnoho, proto bude zobrazen žebříček 5 nejvíce poptávaných a nabízených míst. Poslední část bude zobrazovat srovnání dalších ukazatelů podle krajů jako kriminalitu, nehody nebo nezaměstnanost. Wireframe návržení této obrazovky je na obrázku 5.1.

Dashboard pro kraj

Dashboard pro daný kraj bude mít stejné rozložení jako hlavní dashboard, protože je k dispozici většina stejných ukazatelů. Srovnání krajů bude nahrazeno srovnáním okresů a také žebříček poptávaných a nabízených míst by měl být dostupný i pro okresy.

Dashboard pro zaměstnání

Dashboard pro zaměstnání bude zobrazovat méně informací, protože je k dispozici jen průměrná mzda, počet zaměstnaných osob a poptávka a nabídka po pracovní síle. Poslední část tedy bude věnována srovnání poptávky a nabídky po pracovní síle mezi určitými druhy zaměstnání.



Obrázek 5.1: Wireframe hlavní obrazovky

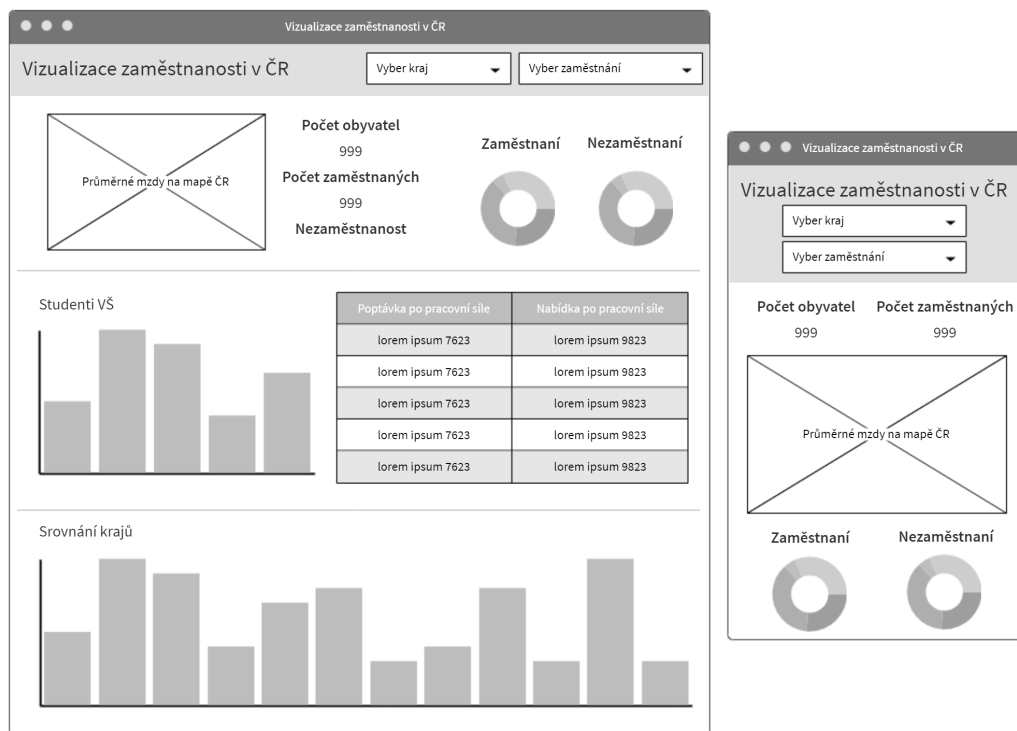
5.3 Responzivita

Responzivní design webových stránek se zabývá optimalizací jejich rozlišení tak, aby odpovídalo velikosti displeje pro všechny druhy nejrůznějších zařízení. V dnešní době hraje responzivní chování webu důležitou roli, například v roce 2018 navštívilo až 54 % uživatelů webové stránky z mobilu nebo tabletu¹.

Ke stylování responzivního webu se využívají flexibilní mřížky, obrázky a vlastnost media queries, která je součástí CSS3 a dokáže rozpoznat velikost displeje. Mobilní verze webu se také snaží o zjednodušování, kdy například skrývají určité prvky stránky nebo celé podstránky [13].

Návrh dashboardu se tedy zabýval i responzivitou, kdy bylo navrženo jednotlivé prvky přeskldávat podle velikosti displeje a některé prvky při menším zobrazení skrýt. Při zobrazení na tabletu by měl dashboard vypadat podobně jako na velkém monitoru a stále by se měl vejít na celou obrazovku. Na mobilní verzi to již není možné, proto by měly být jednotlivé prvky zobrazeny pod sebou viz drátěný model 5.2.

¹Globální statistiky StatCounter viz <http://gs.statcounter.com>



Obrázek 5.2: Wireframe hlavní obrazovky pro zobrazení na tabletu a mobilu

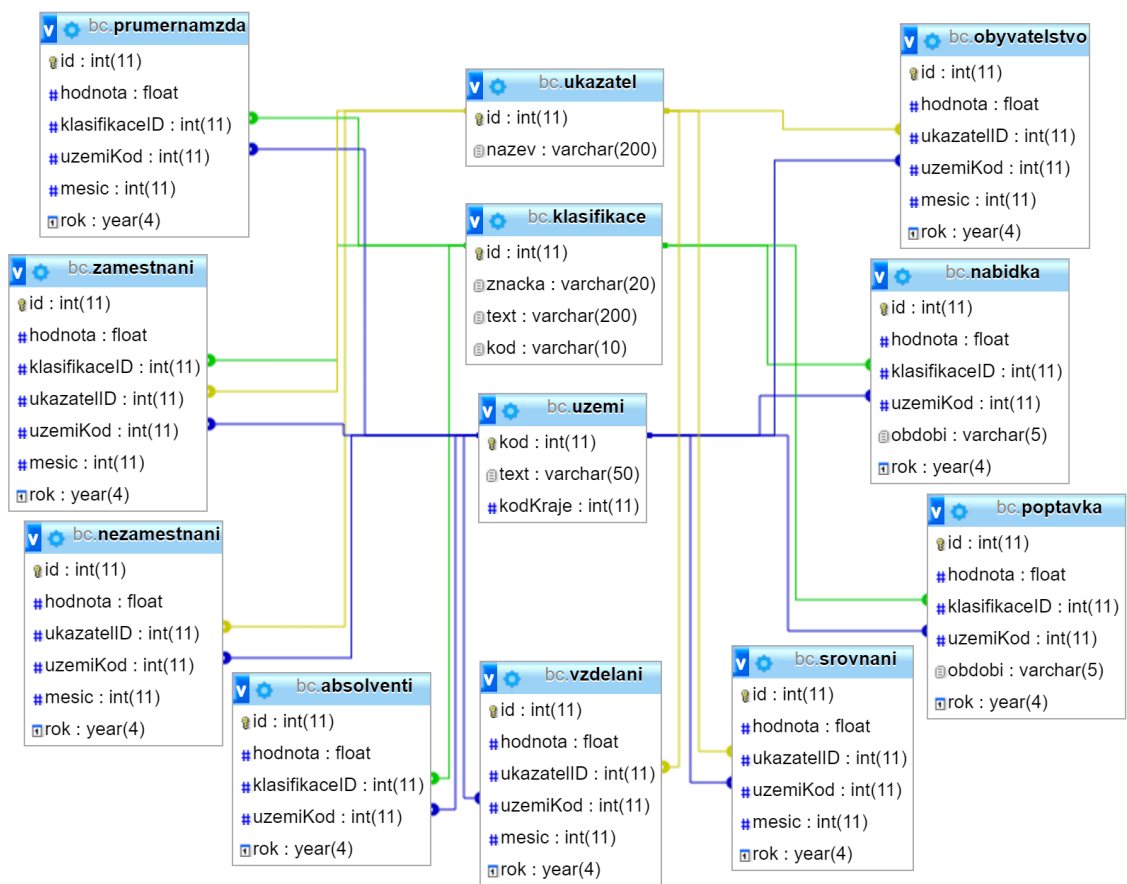
5.4 Návrh databáze

Vizualizovaná data bude potřeba získat ze Statistického úřadu a Ministerstva práce a sociálních věcí. První verze návrhu databáze obsahovala jednu tabulku pro všechny hodnoty, jelikož měly stejné atributy. Nakonec ale byla tabulka s daty rozdělena do několika dílčích tabulek podle jednotlivých ukazatelů, protože pokud by byla všechna data v jedné tabulce, jejich vyhledávání by mohlo být náročné vzhledem k jejich velkému a stále přibývajícimu množství.

Tabulky s hodnotami se odkazují na tabulku s územím, která obsahuje Českou republiku, kraje i okresy. Tabulka s klasifikacemi rozlišuje data podle klasifikace CZ-ISCO nebo podle oborů vzdělání, které jsou využívány pro zobrazení oborů nezaměstnaných absolventů vysokých škol. Poslední pomocnou tabulkou je tabulka ukazatelů, která je využívána, pokud je v tabulce s daty více různých hodnot. Takové tabulky mohly být rozděleny do menších tabulek, ale vzhledem k tomu, že data byla většinou z jedné tabulky Veřejné databáze a budou využívána pro stejné zobrazení, nebyly tyto tabulky rozdělovány. Návrh databáze je zobrazen na obrázku 5.3.

Se stahováním dat také souvisí jejich aktualizace. Většina dat se obnovuje jen jednou za rok, takže by nemělo být obtížné je každý rok ručně aktualizovat. Nicméně některá data se obnovují každý měsíc a v tom případě by měl být vytvořen skript, který data automaticky stáhne, zpracuje a uloží do databáze.

S tím mohou přijít problémy např. v případě změny formátu dat ve Veřejné databázi. Stahování a ukládání dat do databáze by mělo být kompletně oddělené od zbytku systému, aby v případě změny jejich získávání mohlo být jejich zpracování optimalizováno a nenarušilo tak celý zbytek systému.



Obrázek 5.3: Návrh databáze

Kapitola 6

Použité technologie

V této kapitole jsou blíže charakterizovány programovací jazyky a knihovny, které byly pro implementaci informačního systému využity.

6.1 Python

Skriptovací jazyk Python byl použit na stahování dat ve formátu **XML** (z anglického Extensible Markup Language), jejich zpracování a ukládání do databáze. Python je tzv. vícepáradigmatický jazyk, což znamená, že podporuje objektově orientované, procedurální a v menší míře také funkcionální paradigma.

Jeho výhodou je bohatá standardní knihovna, která umožňuje např. stahovat soubory z internetu, rozbalovat komprimované složky nebo vytvořit celý webový server. Python je velmi expresivní, což umožňuje napsat kód na méně řádků ve srovnání s jinými jazyky, také má snadno čitelný kód a je často doporučován začátečníkům [16].

Python také patří mezi nejpobulárnější jazyky. Podle indexu popularity mezi programovacími jazyky nazvaném TIOBE¹ je v květnu 2019 na 4. místě [17]. Navíc podle indexu PYPL² je Python dokonce na prvním místě [14].

6.2 MySQL

S vytvořením databáze a jejím ovládním souvisí systém **MySQL** (z anglického My Structured Query Language), který je nejpobulárnější open source systém řízení báze dat. MySQL pracuje s relačními databázemi a manipulací s daty umožňuje programovací jazyk SQL. Pro snadnější správu relační databáze byl využíván nástroj Adminer, který byl vyvinut jako jednodušší alternativa k phpMyAdmin.

6.3 PHP

PHP: Hypertextový Preprocesor je skriptovací programovací jazyk původně určený právě pro vytváření dynamických webových stránek společně s MySQL, který pracuje na straně

¹Index TIOBE (z anglického The Importance of Being Earnest) je index popularity programovacích jazyků založený na analýze hodnocení odborníků, poptávky po kurzech a požadavků firem.

²Index popularity PYPL (z anglického PopularitY of Programming Language) vzniká analýzou nejvíce vyhledávaných tutoriálů na vyhledávací Google.

serveru. V dnešní době 79 % ze všech webových stránek využívá právě jazyk PHP. Podle indexu popularity TIOBE je aktuálně na 8. místě [17].

Velké změny v programování přinesla verze PHP7, která byla vydána na konci roku 2015. Aktuálně ji využívá již 32 % ze všech stránek založených na PHP a stále přibývají [19]. PHP7 oproti předchozímu PHP5 má dvojnásobnou rychlost, snadněji zpracovává chyby, podporuje deklaraci návratových typů a trojcestný operátor.

Budování složitých webových stránek a aplikací vyžadovalo mnoho času a úsilí, proto vznikly PHP frameworky, které zjednodušují vývoj a nabízí dobře organizovanou strukturu kódu v podobě MVC (z anglického Model View Controller). Také zajišťují bezpečnost webu a prosazují moderní webové vývojové postupy. V současnosti je takových frameworků nespočet a je velmi těžké nějaký zvolit. Mezi tři nejpopulárnější patří Laravel, Symfony a CodeIgniter. Nette framework, který byl následně zvolen, je v současnosti asi na 11. místě, nicméně v roce 2015 byl podle SitePoint na třetím [15].

Nette Framework

Nette framework³ byl spuštěn jako open source v roce 2008. V České republice má velkou komunitu pravděpodobně díky tomu, že jeho autorem je český vývojář David Grudl. Nette nabízí velkou řadu knihoven, umožňuje logování chyb a především se soustředí na bezpečnost webových stránek. Brání se proti Cross Site Scripting (XSS), Cross-Site Request Forgery (CSRF), URL útokům nebo Session hijacking.

Vytvoření projektu v Nette je doporučováno provádět pomocí nástroje Composer, který slouží pro správu závislostí v PHP a umožňuje tak deklarovat pro každý PHP projekt specifické knihovny.

6.4 HTML, CSS, JavaScript

Standardní značkovací jazyk HTML (z anglického Hypertext Markup Language) společně s kaskádovými styly (CSS z anglického Cascading Style Sheets) a jazykem JavaScript tvoří hlavní prostředky pro tvorbu webových stránek a aplikací na klientské straně.

Jazyk HTML obsahuje značky (tzv. *tagy*) a jejich vlastnosti, pomocí nichž definuje strukturu dokumentu. Styly jednotlivých elementů z HTML jsou definovány pomocí jazyka CSS, který se na ně odkazuje pomocí tzv. *selektorů*. Dynamičnost webové stránce dodává právě skriptovací jazyk JavaScript, který je spouštěn až po načtení stránky u klienta.

Níže jsou popsány knihovny, které byly pro tvorbu systému použity a usnadnily tak implementaci klientské části.

Bootstrap

Bootstrap⁴ je open source framework, který byl vydán v roce 2011 a byl vyvinut, aby zjednodušil stylování webové stránky. Poskytuje návrhářské šablony pomocí CSS tříd, které definují styly fontů, tlačítek, formulářů a dalších komponent. Velkou výhodou při tvoření struktury stránky je uspořádání do mřížky o dvanácti sloupcích (tzv. *grid system*), kterou Bootstrap definuje pomocí CSS funkce *flexbox*. Třída *row* definuje řádek a třídy *col-1* až *col-12* určují velikosti sloupců.

³Dostupné na <https://nette.org/>

⁴Dostupné na <https://getbootstrap.com/>

Věnuje se také responzivité stránek, kdy rozděluje CSS třídy podle pěti šířek displeje a definuje je pomocí zkratk, které jsou uvedeny v seznamu:

- **xs**: menší než 576px
- **sm**: větší než 576px
- **md**: větší než 768px
- **lg**: větší než 992px
- **xl**: větší než 1200px

jQuery

jQuery⁵ je nejpobulárnější open source knihovna založená na jazyce JavaScript, která vznikla v roce 2006. Usnadňuje manipulaci s HTML elementy a jejich vyhledáváním, manipulaci s CSS, řešení událostí nebo animací.

Chosen.js

Chosen.js⁶ je open source knihovna, která využívá jQuery a usnadňuje stylování HTML elementu select. Umožňuje do standardního výběru přidat textové vyhledávání, upravuje výběr více položek najednou a umožňuje některé položky zakázat.

Chart.js

Chart.js⁷ je jedna z mnoha open source knihoven jazyka JavaScript, která umožňuje zobrazování grafů na webové stránce. Poskytuje osm druhů grafů:

- sloupcové grafy
- skládané grafy
- spojnicové grafy
- bodové grafy
- plošné grafy
- koláčové grafy
- prstencové grafy
- paprskové grafy

Všechny grafy jsou zobrazovány pomocí HTML elementu *canvas* a jsou responzivní a interaktivní. Chart také umožňuje nastavení více os, velikosti kroků, minima a maxima. Důležitou součástí je nastavování stylů grafů jako např. velikosti fontů, barvy nebo zobrazení pomocných os.

⁵Dostupné na <https://jquery.com/>

⁶Dostupné na <https://harvesthq.github.io/chosen/>

⁷Dostupné na <https://www.chartjs.org/>

Kapitola 7

Implementace

Celý systém byl rozdělen na dvě spolu nepropojené části. První část se stará o stahování a ukládání dat do databáze a druhá část data zpracovává a vytváří výsledný dashboard na webové stránce. Případná změna získávání dat a jejich formátu by se měla projevit právě v první části systému a druhou část by neměla žádným způsobem ovlivnit.

7.1 Zpracování dat

V první řadě bylo potřeba naprogramovat stahování a parsování dat z formátu XML, což zajišťuje program psaný v jazyce Python. Stahování souborů zajišťuje knihovna `requests` a pro práci s XML souborem byla využita knihovna `xml.etree.ElementTree`, která pracuje se souborem jako se stromem (tree) a elementy představují jednotlivé uzly stromu. Nakonec bylo potřeba data uložit do databáze, což bylo umožněno knihovnou `MySQLdb`.

Většina dat je získávána z Veřejné databáze, která umožňuje stažení dat právě ve formátu XML, které je pro každou tabulku stejně strukturované. Pro získání podrobnějších informací o poptávce a nabídce po pracovní síle byla tato data získána z Integrovaného portálu Ministerstva práce a sociálních věcí, kde jsou hodnoty klasifikované podle zaměstnání a podle okresů. Ze stejného zdroje byla ještě stažena data o nezaměstnanosti absolventů vysokých škol, která jsou dostupná pro jednotlivé obory a okresy. Formát těchto dat byl také dostupný v XML.

Část programovaná v jazyce Python obsahuje jen několik souborů. Každý soubor stahuje, zpracovává a ukládá data do databáze pro danou tabulku, pokud v ní takové hodnoty ještě nejsou. Spojení s databází zajišťuje soubor `Database.py`. Aby mohla být data pravidelně aktualizována, je připraven soubor `update.py`, který spustí všechny skripty najednou.

7.2 Informační systém

Druhá část systému je naprogramovaná v PHP frameworku Nette a stará se o serverovou i klientskou část webové stránky. Hlavní strukturu webové stránky definuje šablonovací soubor `@layout.latte`, do kterého Nette později vkládá potřebný obsah. Je v něm popsána HTML hlavička a také menu, které obsahuje nadpis a dva filtry. Tím je zaručena stejná struktura HTML dokumentu pro každou podstránku.

Přípravu dat z databáze pro hlavní obrazovku zajišťuje `HomepagePresenter.php` ve funkci `renderDefault()`. Strukturu HTML obsahu určuje soubor `default.latte` v adresáři `Homepage`. V jeho první části jsou definovány všechny HTML elementy, vloženy některé

hodnoty a vytvářeny tabulky. Druhá část je napsaná v jazyce JavaScript a obsahuje tvorbu dynamických prvků jako jsou grafy a mapa. Data jsou v souborech `.latte` získávána pomocí maker, např. `{ $\$$ variable}`, která také umožňují vytvářet podmíněné příkazy nebo cykly.

Po výběru kraje nebo zaměstnání je uživatel přesměrován na jeho podstránku a pomocí metody *GET* je předán jednoznačný identifikátor. Pro získávání a vytváření obsahu stránky je využit stejný princip jako u hlavní obrazovky.

7.2.1 Prvky obrazovek

V této části je popsána implementace některých prvků na obrazovkách. Jedná se převážně o grafické prvky na hlavní stránce, protože prvky na ostatních stránkách jsou velmi podobné.

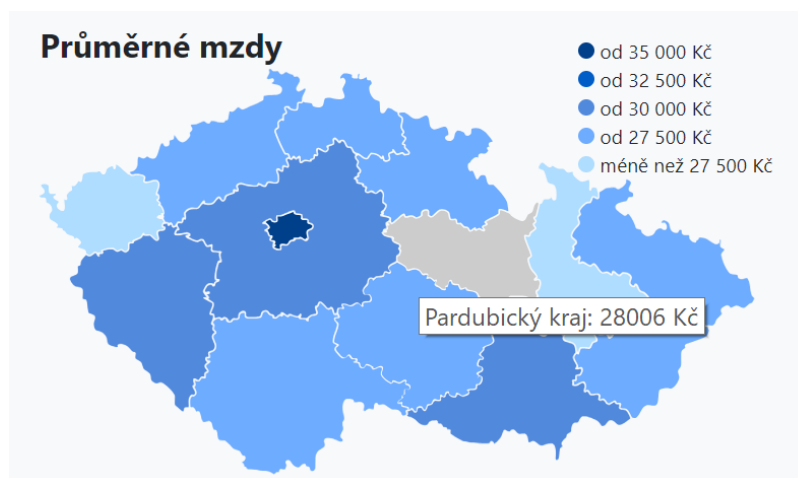
Výběr zaměstnání

Filtr na výběr zaměstnání obsahoval mnoho položek a pro uživatele by nebylo příjemné všechny položky procházet, než by našel tu, kterou hledá. Proto byl výběr upraven pomocí knihovny *Chosen.js*, díky níž může uživatel začít psát název zaměstnání a výběr se mu upraví podle napsané hodnoty.

Mapa ČR

Požadovaná mapa České republiky, kde se měly zobrazovat průměrné platy podle krajů, byla implementována pomocí mapy rozdělené na kraje ve formátu *SVG* a knihovny *jQuery*. Po načtení stránky jsou jednotlivé kraje přebarveny podle dostupných hodnot rozdělených do pěti kategorií. Při najetí myši na daný kraj se zobrazí jeho průměrná mzda (obrázek 7.1) a při kliknutí na kraj se uživatel dostane na obrazovku o daném kraji.

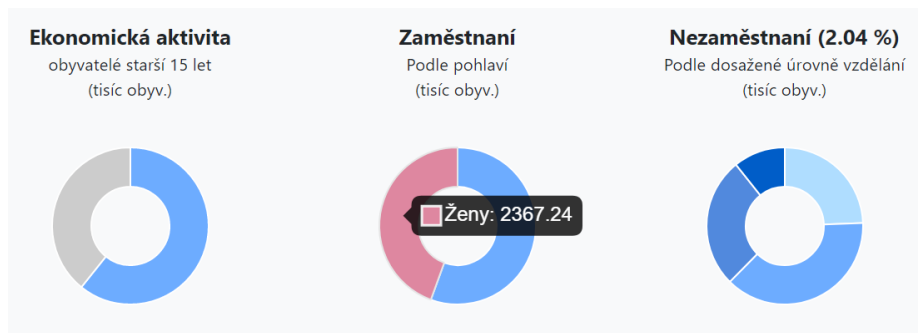
Stejným způsobem je mapa zobrazena i na obrazovce o daném zaměstnání, kde jsou hodnoty pouze pro daný obor. Jelikož se průměrné platy dost liší v závislosti na zaměstnání, stalo se, že například mapa pro pomocné pracovníky byla obarvena pouze do světlé barvy. Jedním řešením bylo automatické vytvoření pěti kategorií v závislosti na hodnotách, ale nakonec nebylo uskutečněno. Pokud tedy uživatel najde mapu celou ve světlé barvě, bude mu jasné, že daný obor má nízké mzdy.



Obrázek 7.1: Průměrné mzdy zobrazené na mapě na hlavní obrazovce

Prstencové grafy

Pro zobrazení prstencových nebo tzv. *doughnut* grafů byly nakonec vybrány tři ukazatele, přičemž ekonomická aktivita obyvatel je v menším rozlišení skryta. Knihovna Chart.js zajišťuje, že při najetí na část grafu se zobrazí její hodnota, jak je vidět na obrázku 7.2.



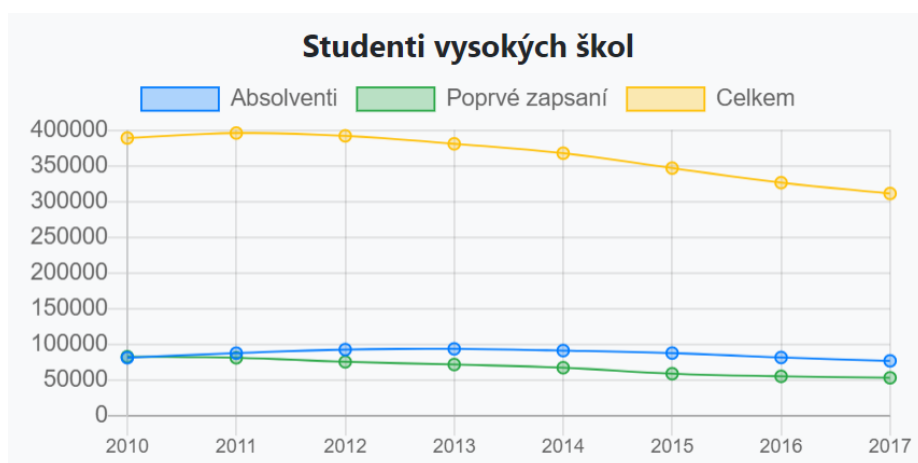
Obrázek 7.2: Příklad prstencových grafů na hlavní obrazovce

Tabulky

Vypsání dat do tabulek bylo poměrně jednoduché díky makru `{foreach}` v souborech `.latte`. Vzhledem k tomu, že na hlavní obrazovce tabulka s poptávkou a nabídkou po pracovní síle zobrazuje jen 5 položek, byla přidána možnost zobrazení celé tabulky v novém okně. Pokud tedy uživatele zajímají kompletní data, má možnost si je prohlédnout.

Studenti VŠ

Spojnicový graf vytvořený pomocí knihovny Chart.js se studenty vysokých škol je zobrazen jen na hlavní obrazovce a zobrazuje počet absolventů, poprvé zapsaných i celkový počet. Aby mohla být tato data určitým způsobem srovnávána, byla zobrazena na časové ose, jak je vidět na obrázku 7.3.

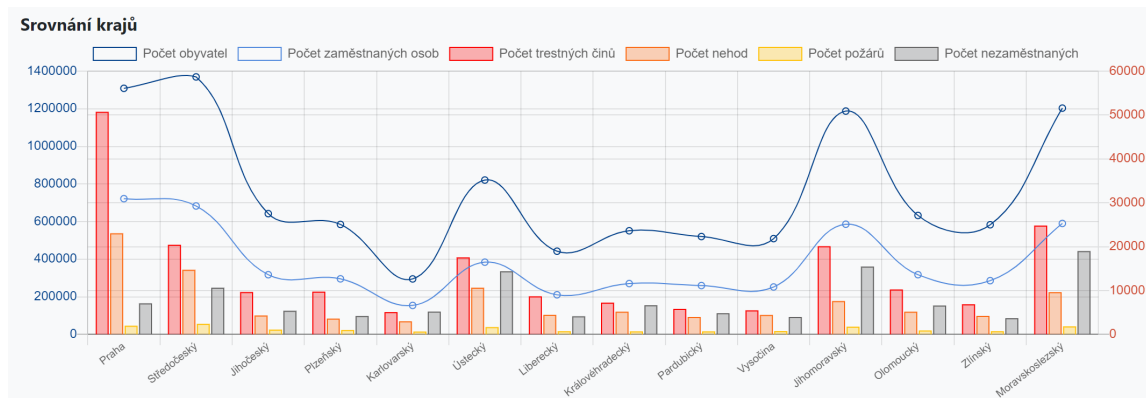


Obrázek 7.3: Výsledný graf zobrazující počty studentů vysokých škol

Srovnání krajů

Graf se srovnáním krajů dominuje spodní části obrazovky. Nejprve zde byly zobrazeny hodnoty týkající se kriminality, nehod a nezaměstnanosti, nicméně pak byly přidány i hodnoty počtu obyvatel a zaměstnaných, aby měl uživatel možnost zařadit si informace do souvislostí. Problém s velkým rozdílem hodnot byla vyřešen dvěma osami y, které jsou odlišeny barevně, jak je vidět na obrázku 7.4. Podobně byl řešen graf se srovnáním okresů zobrazený v obrazovce věnující se kraji.

Zobrazit tento graf na mobilní verzi nebylo možné, proto byl otočen do horizontální polohy a počet obyvatel a zaměstnaných byl zrušen.



Obrázek 7.4: Graf srovnávající vybrané hodnoty krajů na hlavní obrazovce

Výběr okresu

V obrazovce, která se věnuje danému kraji, bylo potřeba vyřešit zobrazování tabulek pro všechny jeho okresy. Výběr okresu je řešen pomocí tzv. *radio button* HTML elementu a v závislosti na jeho vybrání se dynamicky mění jen část obrazovky (viz obrázek 7.5). Tato funkcionality byla řešena pomocí AJAX požadavku, kdy byla část dokumentu s přepínači a tabulkami obalena do formuláře makrem `{form}` a měnící se tabulky obaleny makrem `{snippet}`. Překreslení tabulek řeší funkce `handleOkresy()` v `RegionPresenter.php`.

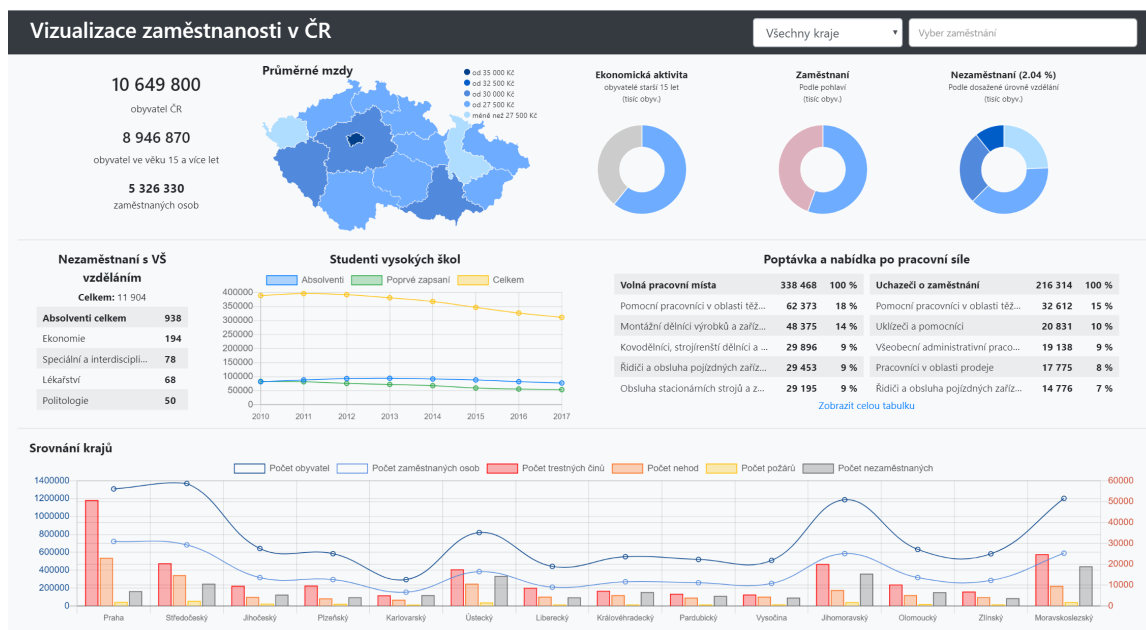
Okres	Nezaměstnaní absolventi		Poptávka a nabídka po pracovní síle					
<input checked="" type="radio"/> Blansko	Celkem	10	Volná pracovní místa	979	100 %	Uchazeči o zaměstnání	1 900	100 %
<input type="radio"/> Brno-město	Ekonomie	5	Kovodělníci, strojřeniští děl...	199	20 %	Pomocní pracovníci v oblas...	378	20 %
<input type="radio"/> Brno-venkov	Elektrotechnika, teleko...	1	Pomocní pracovníci v oblas...	139	14 %	Všeobecní administrativní ...	172	9 %
<input type="radio"/> Břeclav	Lékařství	1	Řidiči a obsluha pojiždných...	89	9 %	Pracovníci v oblasti ochran...	127	7 %
<input type="radio"/> Hodonín	Geografické obory	1	Řemeslníci a kvalifikovaní p...	77	8 %	Pracovníci v oblasti prodeje	113	6 %
<input type="radio"/> Vyškov	Politologie	1	Zpracovatelé potravin, dřev...	65	7 %	Řemeslníci a kvalifikovaní p...	112	6 %
<input type="radio"/> Znojmo								

[Zobrazit celou tabulku](#)

Obrázek 7.5: Část obrazovky umožňující výběr okresu v Jihomoravském kraji

7.2.2 Rozložení prvků

Struktura obrazovky byla implementována pomocí mřížky poskytované knihovnou Bootstrap. Všechny prvky mají definované šířky pomocí CSS tříd col-1 až col-12 a jsou skládány vedle sebe. Aby bylo zajištěné responzivní zobrazení, má každý prvek tyto šířky definované pro jednotlivé velikosti displeje. Drobné úpravy jsou také prováděny v souboru style.css pomocí funkce media queries. Výsledné rozložení prvků na hlavní obrazovce je možné vidět na obrázku 7.6.



Obrázek 7.6: Výsledné rozložení prvků na hlavní obrazovce

Kapitola 8

Testování

Aby byl systém uživatelsky příjemný a uspokojoval všechny jejich potřeby, je nutné jej testovat. V této kapitole je blíže popsáno uživatelské testování systému a zhodnocení jeho výsledky. Systém s vizualizací dat o zaměstnanosti ČR byl testován pomocí testu použitelnosti, což je jedna z nepoužívanějších metod.

8.1 Průběžné testování

Nejprve byl vytvořen drátěný model, díky němuž jsem se mohla ptát uživatelů, zda jim rozložení prvků připadá srozumitelné. Z jejich odpovědí nebyly zřejmé žádné větší problémy. Ocenili mapu České republiky na viditelném místě a využití celého prostoru stránky. Jelikož ještě neměli možnost si vyhledávání dat vyzkoušet, nebyli schopni najít více problémů.

Další testování bylo provedeno až po naprogramování velké části systému, aby si uživatel mohl vyzkoušet vyhledávání dat. První verzi systému vyzkoušeli dva uživatelé, přičemž odhalili základní problémy. Na základě jejich připomínek bylo upraveno rozložení prvků, přidáno více hodnot v procentuálním vyjádření a přibyla možnost zobrazení celé tabulky uchazečů o zaměstnání a volných míst. Také byly do grafů se srovnáním krajů a okresů přidány další hodnoty.

8.2 Test použitelnosti

Výsledný informační systém byl testován pěti dalšími lidmi, z nichž dva byli studenti a tři zaměstnanci. Během testu byli uživatelé pozorováni, jak postupují při hledání odpovědí na předem připravené otázky, a byly sledovány jejich reakce a připomínky. Mezi otázky, na které uživatelé hledali odpovědi, patřily:

1. Jaká je průměrná mzda v Jihočeském kraji?
2. Převyšuje v poslední době počet poprvé zapsaných studentů absolventy vysokých škol?
3. Který obor studovalo nejvíce nezaměstnaných absolventů?
4. Ve kterém kraji je pravděpodobně nejmenší kriminalita v závislosti na jeho počtu obyvatel?
5. Kolik obyvatel má okres Kroměříž ve Zlínském kraji?

6. Jaká je míra nezaměstnanosti v Jihomoravském kraji?
7. Jaká je průměrná mzda techniků v oblasti informačních a komunikačních technologií?
8. V jaké skupině zaměstnání najdeme nejvíce uchazečů?
9. Ve kterém okrese je pro tuto skupinu zaměstnání nejvíce volných míst?

Větší polovina otázek uživatelům nedělala problém a odpověď hledali na správných místech a dokázali se na stránce zorientovat. Problém často způsobovala otázka č. 4, která byla složitější a zároveň vyžadovala větší zkoumání grafu se srovnáním krajů ve spodní části obrazovky, nicméně i na tu uživatelé po delší chvíli dokázali odpovědět.

Zajímavým zjištěním bylo, že téměř nikdy nepoužívali filtry v pravé horní části stránky. Na obrazovku krajů se intuitivně dostali kliknutím na mapu a proto si připravených filtrů nevěšili. Otázka č. 7 jim proto dělala velké problémy, protože toto zaměstnání nemohli nikde najít. Většinou jej hledali v tabulce volných míst a počtu uchazečů o zaměstnání, kterou si následně zobrazili celou a požadované zaměstnání našli v seznamu. Filtr využil jen jeden uživatel.

8.3 Zhodnocení testů

Z výsledků testů a připomínek testovaných uživatelů můžeme odvodit následující závěry. Rozložení prvků na hlavní obrazovce uživatelům nezpůsobilo velké problémy a i když zpočátku jim chvíli trvalo, než našli ten správný graf, na konci testování se již v systému docela orientovali a odpovědi na poslední otázky nacházeli velmi rychle.

Problém s filtrem na zaměstnání byl pravděpodobně kvůli jeho jinému vzhledu než filtru krajů, ale také mohl být způsoben tím, že prvků na hlavní obrazovce je mnoho a filtr v pravém horním rohu nebyl hned zaznamenán. Bylo by dobré jej zobrazit většími písmeny a možná přidat ikonu lupy, která uživateli napoví, že zde může hledat.

Další poznatek byl získán o obrazovce kraje, kde uživatelé intuitivně hledali počet obyvatel v okrese v prostřední části, která umožňuje přepínání mezi okresy a zobrazuje jen nezaměstnané absolventy a poptávku a nabídku pracovní síly. Informace o počtu obyvatel, počtu zaměstnaných a průměrné mzdě by zde proto mohly být také zobrazeny.

Uživatelé opět oceňovali a často používali mapu s možností prokliku na kraj a podobnou funkčnost hledali při kliknutí na název kraje v grafu se srovnáním. Někteří uživatelé by uvítali více vysvětlivek k hodnotám a také více interaktivních možností jako například zobrazování dalších hodnot v grafu se srovnáním krajů.

Kapitola 9

Závěr

Cílem této práce bylo vytvoření informačního systému, který měl vhodně vizualizovat statistická data z oblastí zaměstnanosti, nezaměstnanosti, průměrných mezd a uplatnění studentů vysokých škol. K tomuto účelu byl zvolen vizualizační nástroj dashboard.

Nejprve byla charakterizována data, která poskytuje Český statistický úřad, a následně byly podrobně popsány vybrané oblasti dat. V další části práce byl představen nástroj dashboard, principy jeho zobrazení a tvorby. Následně byl proveden průzkum uživatelů a bylo zjištěno, které informace je zajímají.

Po provedení analýzy bylo možné vytvořit první návrhy prvků obrazovek a jejich rozložení na stránce, po čemž následovalo navržení databáze. Následně byly naprogramovány skripty, které ji naplnily potřebnými daty a zároveň umožnily jejich budoucí aktualizaci. Poslední důležitou fází byla samotná implementace informačního systému, která vyžadovala vytvoření jednotlivých prvků a jejich uspořádání na stránce. Nedílnou součástí vývoje bylo také uživatelské testování.

Vytvořený dashboard by měl v budoucnu sloužit převážně studentům, kteří rozhodují o svém povolání a zjišťují si informace o jejich vysněném oboru. Také by měl poskytovat určitý přehled o zaměstnanosti v České republice, díky kterému by uživatel mohl najít nové informace. Systém v současnosti běží na neplacené webové adrese¹.

¹<http://statisticka-data.8u.cz/>

Literatura

- [1] Státní správa. [Online; navštíveno 28.01.2019].
URL <https://www.statnisprava.cz/>
- [2] Český statistický úřad. [Online; navštíveno 28.01.2019].
URL <https://www.czso.cz>
- [3] Boček, J.; Mráček, J., Jakub a Mynarz: *Otevřená data: příležitost pro Českou republiku*. Nadace Open Society Fund, 2012, ISBN 978-80-87725-03-0.
- [4] Eckerson, W. W.: *Performance dashboards: measuring, monitoring, and managing your business*. John Wiley & Sons, 2006, ISBN 978-0-471-72417-9.
- [5] Few, S.: *Information Dashboard Design: The Effective Visual Communication of Data*. O'Reilly, 2006, ISBN 0-596-10016-7.
- [6] Hendl, J.: *Kvalitativní výzkum: základní metody a aplikace*. Portál, 2005, ISBN 80-7367-040-2.
- [7] Hynek, J.: *Informační dashboardy: Význam a použití informačních dashboardů*. 2014.
URL <http://www.fit.vutbr.cz/study/courses/TJD/public/1314TJD-Hynek.pdf>
- [8] Inc., J.: *A guide to creating dashboards people love to use*. 2009.
URL https://static1.squarespace.com/static/52f42657e4b0b3416ff6b831/t/5310292ce4b08d35a87c9426/1393568044420/Guide_to_Dashboard_Design.pdf
- [9] Jacobs, J.; Rudis, B.: *Data, Driven Security: Analysis, Visualization and Dashboards*. John Wiley & Sons, 2014, ISBN 978-1-118-79372-5.
- [10] Johnson, J.: *Designing with the Mind in Mind*. Morgan Kaufmann, 2010, ISBN 978-0-12-375030-3.
- [11] Krug, S.: *Web design: nenutte uživatele přemýšlet!* Computer Press, 2006, ISBN 80-251-1291-8.
- [12] Malik, S.: *Enterprise dashboards: design and best practices for IT*. John Wiley & Sons, 2005, ISBN 978-0-471-73806-0.
- [13] Marcotte, E.: *Responsive Web Design*. A Book Apart, 2011, ISBN 978-1937557188.
- [14] PYPL: PopularitY of Programming Language. [Online; navštíveno 08.05.2019].
URL <http://pypl.github.io/PYPL.html>

- [15] Sitepoint: The Best PHP Framework for 2015. [Online; navštíveno 08.05.2019].
URL <https://www.sitepoint.com/css-optimization-boosting-pwa-performance/>
- [16] Summerfield, M.: *Programming in Python 3: A Complete Introduction to the Python Language*. Addison-Wesley Professional, 2010, ISBN 978-0-321-68056-3.
- [17] TIOBE: Index for May 2019. [Online; navštíveno 08.05.2019].
URL <https://www.tiobe.com/tiobe-index/>
- [18] Tufte, E. R.: *The visual display of quantitative information*. Graphics press Cheshire, CT, 2001, ISBN 978-1-930824-13-3.
- [19] W3Techs: Usage statistics and market share of PHP for websites. [Online; navštíveno 08.05.2019].
URL <https://w3techs.com/technologies/details/pl-php/all/all>
- [20] Český statistický úřad: *Statistický metainformační systém*. [Online; navštíveno 28.01.2019].
URL <http://apl.czso.cz/iSMS/home.jsp>

Příloha A


Obsah CD

- **src/** - zdrojové soubory praktické části
- **doc/** - zdrojové soubory k technické práci
- **BP.pdf** - technická práce ve formátu PDF

Příloha B

Další vytvořená persona

Marek Polčák



Věk: 23
Práce: Student vysoké školy, obor pedagogika, tělocvika, zeměpis
Rodina: Svobodný
Adresa: Brno, ČR

Ambiciózní Extrovert Sportovec

Životopis

Marek studuje poslední ročník vysoké školy a přemýšlí, co bude dělat, až dostuduje. Učitelem se stát nechce, proto zkoumá jiné možnosti uplatnění v praxi.
Ve volném čase se nejvíce věnuje oblíbenému volejbalu, který hraje i závodně, ale také cyklistice a běhání.


Osobnost

Introvert	Extrovert
Myšlení	Cit
Vnímání	Intuice
Pasivní	Aktivní

Informační prostředky

Mobil
Notebook
Televize

Oblíbené aplikace



Motivace

Pobídnutí
Strach
Růst
Síla
Společnost

Cíle

- Dostat se vysoko
- Sportovat
- Poznávat svět

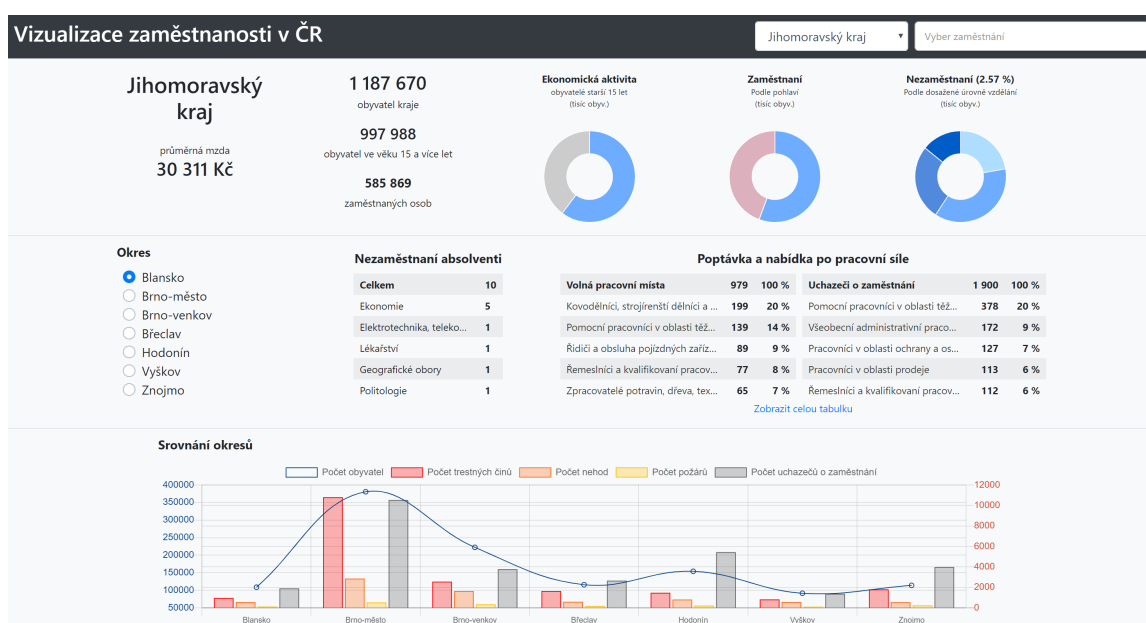
Obavy

- Strach z prohry
- Strach z neúspěchu
- Nudná práce

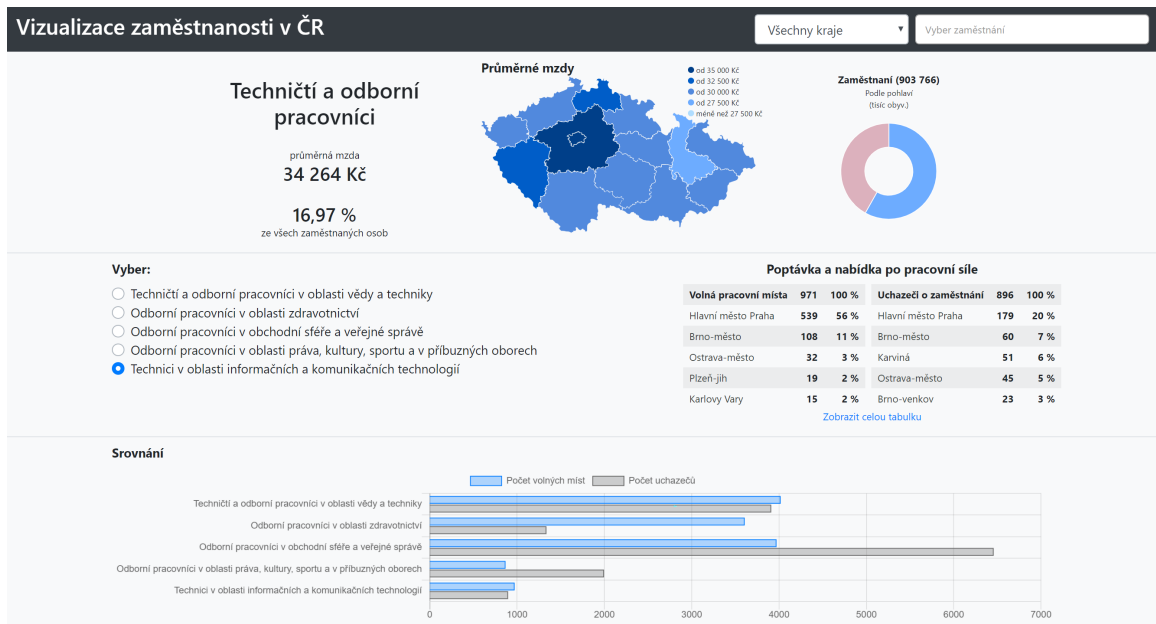
„Žádný zápas nesmíte vzdát. A to ani když zrovna prohráváte.“

Příloha C

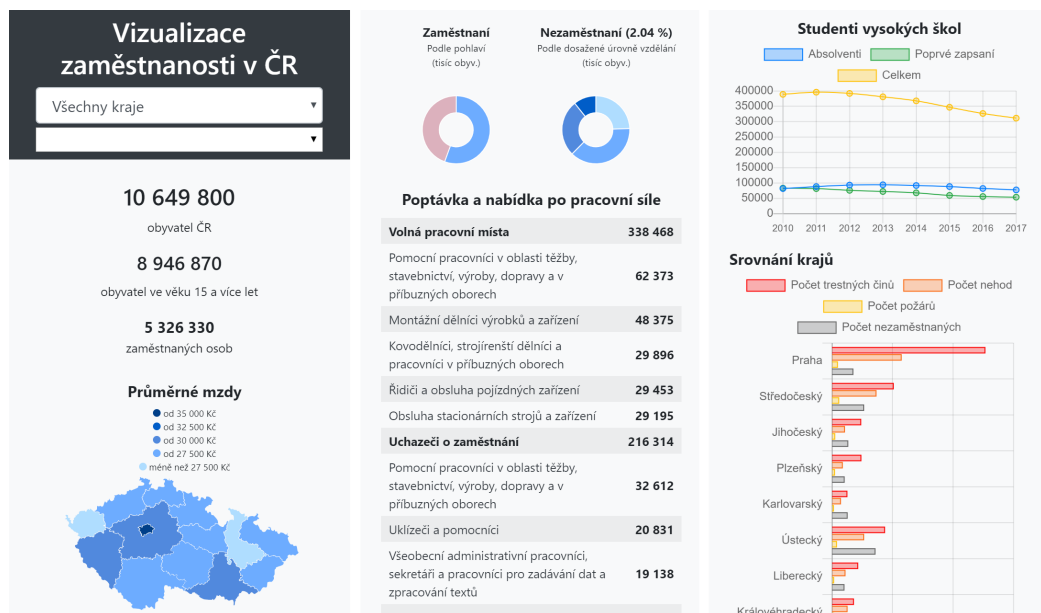
Snímky obrazovek



Obrázek C.1: Příklad obrazovky pro kraj



Obrázek C.2: Příklad obrazovky pro zaměstnání



Obrázek C.3: Příklad hlavní obrazovky zobrazené na mobilu

Příloha D

Zkratky

CSS	Cascading Style Sheets
CSRF	Cross-Site Request Forgery
CZ-NACE	Klasifikace ekonomických činností
CZ-ISCO	Klasifikace zaměstnání
ČSÚ	Český statistický úřad
HTML	Hypertext Markup Language
MPSV	Ministerstvo práce a sociálních věcí
MVC	Model View Controller
MySQL	My Structured Query Language
OSF	Open Society Fund
UI	User Interface
VŠPS	Výběrové šetření pracovních sil
XML	Extensible Markup Language
XSS	Cross Site Scripting