

Univerzita Hradec Králové
Fakulta informatiky a managementu
Katedra informatiky a kvantitativních metod

Nástroje vizualizace dat pro podporu rozhodování

Bakalářská práce

Autor: Daniel Ott
Studijní obor: Informační management

Vedoucí práce: prof. RNDr. Hana Skalská, CSc.

Prohlášení:

Prohlašuji, že jsem bakalářskou práci zpracoval samostatně a s použitím uvedené literatury.

V Hradci Králové dne 31.8.2022

Daniel Ott

Poděkování:

Děkuji vedoucí bakalářské práce prof. RNDr. Hana Skalské, CSc. za metodické vedení práce a pomoc při řešení problémů.

Anotace

Zaměřením této bakalářské práce je oblast vizualizace dat. Práce je zaměřena na programy a nástroje používané pro vizualizaci a vysvětluje klíčové pojmy z této oblasti. Představené jsou programy pro datovou vizualizaci a jejich potenciální využití ve skladovém hospodářství a administrativě středně velké firmy. Data jsou vizualizována pomocí nástrojů Microsoft Power BI, Microsoft Excel a Google Data Studio.

Tato práce je rozdělená na tři části. První část vysvětluje pojmy jako data, vizualizace dat, historii vizualizace dat, důvody a výhody používání vizualizace. Jsou uvedeny zásady pro vytvoření dobré vizualizace a také potenciální nebezpečí špatně zpracované datové vizualizace a zavádějících statistik, práce se v této části také zaměřuje na využití augmentované reality pro datovou vizualizaci. Ve třetí části jsou uvedeny vybrané technologie používané pro vizualizaci dat a jejich možné použití při praktické činnosti ve skladovém hospodářství, personální agendě a kontrole financí.

Klíčová slova

Data, Zdrojová data, Vizualizace dat, Graf, Diagram, Kontingenční tabulka, MS Power BI, MS Excel, Google Data Studio

Annotation

Data Visualization for decision support

The focus of this bachelor thesis is the field of data visualization. The work is focused on programs and tools used for visualization. These explains key concepts from this area. Programs for data visualization and their potential use in warehouse management and administration of a medium-sized company are presented. Data is visualized using Microsoft Power BI, Microsoft Excel and Google Data Studio.

This thesis is divided into three parts. The first part explains concepts like data, data visualization, history of data visualization, reasons and benefits of using visualization. Principles for creating a good visualization are stated, as well as the potential dangers of poorly processed data visualization and misleading statistics. In this part the work is also focusing on the augmented reality application in data visualization. In the third part selected technologies used for data visualization are presented with their possible application in practical warehouse management activities, human resources and financial control.

Keywords

Data, Raw Data, Data visualization, Graph, Diagram, Pivot table, MS Power BI, MS Excel, Google Data Studio

Obsah

1	Úvod	1
2	Cíl práce a metodika.....	2
3	Literární rešerše.....	3
3.1	Základní pojmy vizualizace dat.....	3
3.1.1	Data a zdrojová data	3
3.1.2	Dělení dat do skupin	3
3.1.3	Vizualizace dat	5
3.2	Důvody, proč vizualizovat data.....	6
3.3	Historie vizualizace dat.....	7
3.4	Využití rozšířené reality pro vizualizaci dat.....	12
3.4.1	Praktické využití rozšířené reality ve ŠKODA AUTO	13
4	Volba způsobů řešení	14
4.1	Jak správně vizualizovat data	14
4.1.1	Metriky vizualizace	15
4.2	Nebezpečí zavádějící vizualizace dat.....	15
4.3	Schémata, grafy a diagramy	17
4.3.1	Schéma	17
4.3.2	Grafy a diagramy	17
4.3.3	Kontingenční tabulka	27
5	Technologie vizualizace dat.....	28
5.1	Microsoft Excel.....	28
5.2	Microsoft Power BI	30
5.3	Google Data Studio.....	33
5.4	Další nástroje pro vizualizaci dat.....	35

5.4.1	Jazyk R.....	37
6	Využití vizualizace dat ve firemním prostředí.....	38
6.1	Vizualizace dat finanční administrativy za použití Microsoft Excel	39
6.1.1	Podkladová data a postup zpracování.....	40
6.1.2	Výsledné vizualizace a závěry.....	41
6.2	Vizualizace dat personální agendy za použití Google Data studio ...	47
6.2.1	Podkladová data a postup zpracování.....	47
6.2.2	Výsledné vizualizace a závěry.....	48
6.3	Vizualizace dat skladového hospodářství za použití Microsoft Power BI Desktop	53
6.3.1	Podkladová data a postup zpracování.....	53
6.3.2	Výsledné vizualizace a závěry.....	54
7	Shrnutí výsledků	67
8	Závěry a doporučení	68
9	Seznam použité literatury	69
10	Přílohy	72

Seznam obrázků

Obrázek 1 – Mapa oblasti zasažené nákazou od Johna Snowa [9]	10
Obrázek 2 – Minardova mapa Napoleonova ruského tažení [9]	11
Obrázek 3 – zavádějící statistika Fox News [18]	16
Obrázek 4 – zavádějící schéma české vlády [16]	17
Obrázek 5 – Korelační diagram a Histogram [20]	19
Obrázek 6 – Liniový, Spline a Plošný graf, Regulační diagram [20]	20
Obrázek 7 – Trendline a Svícový graf [20]	21
Obrázek 8 – Kruhový diagram, Víceúrovňový kruhový diagram a Polar Area Graf [20]	22
Obrázek 9 – Stromová mapa exportu firem v ČR [21]	23
Obrázek 10 – Populační pyramida a Solid Gauge Chart [20]	24
Obrázek 11 – Word Cloud a Cyklický diagram [20]	24
Obrázek 12 – Profilová mapa a topografická mapa [20]	25
Obrázek 13 – Kartogram světové populace [20]	26
Obrázek 14 – SWOT Analýza a Teplotní mapa [20]	27
Obrázek 15 – Pracovní plocha aplikace MS Excel 2016	29
Obrázek 16 – Pracovní plocha aplikace Microsoft Power BI Desktop	32
Obrázek 17 – Interaktivní graf vykreslený aplikací AnyChart [27]	36
Obrázek 18 – 3D graf vykreslený aplikací Walrus [28]	37
Obrázek 19 – bodový graf vytvořený pomocí knihovny ggplot2 [30]	38
Obrázek 20 – Použitý postup úpravy dat a tvorby diagramu	41

Seznam schémat

Schéma 1 – Kdo nejvíc nakupuje a utrácí.....	42
Schéma 2 – Složení celkových výdajů za období leden-červen	43
Schéma 3 – Nejčastěji nakupované zboží v kanceláři.....	44
Schéma 4 – Proplacené a uskutečněné nákupy	45
Schéma 5 – Kontrola financí v čase a dodržování limitu.....	45
Schéma 6 – Počet vyhovujících řidičů	49
Schéma 7 – Druhy úvazků zaměstnanců	50
Schéma 8 – Dojezdová vzdálenost zaměstnanců do práce.....	51
Schéma 9 – Zastoupení žen a mužů mezi zaměstnanci	52
Schéma 10 – Typy a výrobci GPU na skladě.....	55
Schéma 11 – GPU z haly 1 v servisu	56
Schéma 12 – GPU z haly 2 v servisu a podmínka filtru v Power BI.....	57
Schéma 13 – RAM a CPU určené k osazení	58
Schéma 14 – Základní desky určené k osazení	59
Schéma 15 – Otestované a osazené sestavy	60
Schéma 16 – Výchozí data k další tvorbě sestav	61
Schéma 17 – Nefunkční a poškozené komponenty	62

Seznam tabulek

Tabulka 1 – Přehled financí v kanceláři za období leden-červen	46
Tabulka 2 – Data zaměstnanců pobočky	52
Tabulka 3 – Všechny typy a výrobci karet ve skladu.....	63
Tabulka 4 – Typy GPU z haly 1 v servisu	64
Tabulka 5 – Typy GPU z haly 2 v servisu.....	64
Tabulka 6 – RAM k osazování sestav	65
Tabulka 7 – CPU k osazování sestav.....	65
Tabulka 8 – Základní desky pro sestavy.....	65
Tabulka 9 – Kompletní sestavy	65
Tabulka 10 – Záznamy o poškozených a nefunkčních komponentách.....	66

1 Úvod

Bakalářská práce se věnuje vizualizaci dat a v rámci toho jsou v ní uvedeny příklady grafů, diagramů a schémat a také zásady tvorby kvalitních vizualizací. Vizualizace dat z firemního prostředí, jako jsou finanční výkazy, data zaměstnanců či stav produktů ve skladu, je vhodná pro koordinaci rozhodování o pracovních procesech.

V každodenním životě vnímá člověk obrovské množství informací. a naprostá většina těchto informací je vnímána zrakem. Mozek proto dokáže daleko lépe porozumět informacím podaným vizuálně. Pokud se pracuje s velkými objemy dat jsou vizualizace pro pochopení souvislostí nezbytné. Informace vnímané vizuálně jsou snazší na zapamatování a pochopení. Proto se zejména při prezentování, výuce, a poslední dobou vzhledem k všudypřítomným obrazovkám i v běžné konverzaci, snažíme využívat diagramů, grafů a jiných forem vizuálního uspořádání dat.

Tato práce představí nejrozšířenější a nejpoužívanější formy a způsoby vizualizace dat z různých oblastí a také technologie, které se pro vizualizaci dat používají.

2 Cíl práce a metodika

Cílem této bakalářské práce je za pomoci prezentovaných technologií představit sféru vizualizace dat a navrhnout její využití při řízení činností v prostředí středně velké firmy. Autor předá poznatky z oblasti vizualizace a tím vysvětlí nejdůležitější pojmy, které s touto oblastí souvisí. Součástí práce je analýza dat skladového hospodářství a ze závěrů analýzy budou učiněna rozhodnutí o změnách v pracovních procesech. Nejprve budou popsány vybrané technologie a následně za jejich použití proběhne analýza výsledků a budou prezentována řešení aktuální nepříznivé situace ohledně skladového hospodářství.

Autor vychází z odborné literatury věnující se datům, vizualizaci dat a historickému kontextu vizualizace dat. V teoretické části práce vychází také z informací a dokumentů publikovaných společnostmi, které se věnují oblasti Business Intelligence (BI) a vývoji softwaru pro vizualizaci podnikových dat. Postupy využití při tvorbě praktické části se řídí zásadami pro tvorbu dobrých vizualizací uvedených v teoretické části. Uvedené vizualizační prostředky (schémata, grafy, diagramy) jsou nejprve popsány obecně a následně jsou u některých z nich prezentovány praktické aplikace. U převzatých schémat jsou uvedeny zdroje.

Schémata zpracovaná autorem vychází z dat poskytnutých společností zabývajících se údržbou a logistikou počítačových komponent. Závěry této práce budou aplikovány v činnosti firmy, zejména závěry vyvozené z analýzy dat skladového hospodářství. Ke zpracování dat jsou využity nástroje Microsoft Excel, Microsoft Power BI a Google data studio.

3 Literární rešerše

V této části se bude práce zabývat samotnou vizualizací dat a vysvětlí také za pomoci uvedených zdrojů klíčové pojmy z oblasti vizualizace. Také bude vysvětleno, proč je vhodné vizualizovat data a představen bude historický vývoj vizualizace dat. V poslední kapitole se práce věnuje technologii rozšířené reality a její aplikaci v oblasti vizualizace dat.

3.1 Základní pojmy vizualizace dat

3.1.1 Data a zdrojová data

Z pohledu výpočetní techniky by se jako data daly definovat jako množiny informací, které byly zpracovány tak, že práce a manipulace s nimi je jednodušší než s takzvanými Raw Data (někdy též nazývána zdrojová, primární nebo atomická data.) To jsou data, která pocházejí z prvotního pozorování zdroje a doposud s nimi nebyly provedeny operace jako „čištění“ od odchylek nebo nevalidních a chybných hodnot. Nejčastěji se jedná o prosté číselné hodnoty reprezentující a popisující určitou skutečnost a objekty reálného světa, například fyzikální veličiny. [1]

Raw Data se vznikají dvěma způsoby. Zaprvé cíleným sběrem a pozorováním, potom se označují jako Captured data. Naproti tomu data, která se označují jako Exhaust data jsou vedlejším produktem činnosti nějakého systému, která nepřímo souvisí s primární funkcí. Běžný příklad z praxe je pokladna v obchodě, která kromě primární funkce – kalkulace ceny – může zachytávat data jako číslo slevové karty a typy nakupovaných produktů a na základě toho pak obchod personalizuje akce pro zákazníka. [2]

3.1.2 Dělení dat do skupin

Pokud budeme vycházet z prací Efektivní analýza a využití dat (autoři: Zach Gemignani a kolektiv) a Vizualizace více rozměrných dat (autor: Jaroslav Myslivec), můžeme data dělit na několik skupin. Jeden pohled uvádí dvě hlavní skupiny dat:

Kvantitativní typ dat – tato skupina dat je charakteristická tím, že je vyjádřena číselnou hodnotou a v takovém případě je možné s nimi provádět například aritmetické operace anebo na ně aplikovat matematické funkce.

Kategoriální typ dat – v tomto případě číselná hodnota neodpovídá přímé reprezentaci daných dat, toto číslo je v takovém případě spíš odkaz. Je to kódové označení určitého nepočítatelného jevu (hexadecimální kód barvy), kterým se označuje daná kategorie a nelze na ně aplikovat matematické funkce ani aritmetické operace.

Naopak pro obě skupiny je možné v závislosti na použité technologii použít funkce jako vyhledávání nebo filtraci a rovněž nástroje pro manipulaci s daty (editace, mazání, kopírování...) [3][4][5]

Z pohledu statistiky se dají data rozdělit na jiné skupiny, a sice [3][4][5]:

Data nominální – tato data se také označují jako kvalitativní. Význam jejich hodnot nejde porovnat, nelze je nijak postupně seřadit, protože u nich neexistuje nic jako velikost. Ve statistickém souboru na ně lze tedy aplikovat například pouze funkce jako četnost, nebo z nich lze určit modus. Takzvaná dichotomická data jsou podskupinou nominálních dat, nabývají pouze logických hodnot jako ANO / NE nebo 1 / 0.

Data ordinální – nabývají jednoduchých číselných hodnot a oproti nominálním datům u nich tedy lze zavést přirozené uspořádání. Mezi dvěma různými hodnotami je možné jasně určit, která z nich je větší a která menší.

Data kvantitativní – Lze je měřit na číselné stupnici, jedná se o prosté číselné veličiny.

Data intervalová – v intervalových datech má kromě porovnávání hodnot také smysl hodnotit i vzdálenosti mezi jednotlivými nebo hodnotami nebo kategoriemi hodnot

Data poměrová – též nazývána podílovými daty. Představují taková data, u kterých jsou již pevně definovány i poměry mezi hodnotami. Patří sem například všechny fyzikální veličiny definované v souladu se standardem SI. Také mají tato data jasně

definovanou absolutní nulu, tj. nejnižší možnou hodnotu, kterou lze zachytit a mohou mít určenou i maximální hodnotu.

Data poměrová a intervalová jsou obvykle soubory spojitých veličin a jejich hodnoty se mohou plynule měnit v průběhu určitého intervalu. Naproti tomu data ordinální a data nominální jsou diskrétní, to znamená, že mohou nabývat jen hodnot, které byly předem definovány a mají konečný počet. Často se spojitá data převádí na diskrétní. Obvyklým způsobem je prosté seskupení jednotlivých hodnot do intervalů, každý interval pak představuje jednu kategorii, která je v diskrétní formě kvůli dalšímu zjednodušení práce s daty. [3][4][5]

3.1.3 Vizualizace dat

Podle veřejně dostupného slovníku Merriam-Webster se dá vizualizace definovat jako: 1) vytváření mentálních vizuálních představ, 2) činnost a proces tlumočení ve vizuálních termínech nebo uvedení do viditelné formy.

Pro pojem Vizualizace dat existuje velké množství definic a ty se ve svém znění někdy rozcházejí. Jedna z nejužitečnějších českých definic zní takto: "*Vizualizace dat je proces zkoumání dat a informací po jejich převedení do grafické podoby. Jejím cílem je pochopení zkoumaných jevů a vniknutí do problému. Proto o vizualizaci mluvíme též jako o vizuální analýze dat.*" [6].

Vizualizace dat zjednodušuje vnímání obsahu dat. Právě proto jsme v našem okolí neustále vystaveni datům ve vizuální podobě. V některých případech by se dalo říct, že vizualizace dat je pro určité lidi nutná k jejich práci, zejména v odvětvích jako finance nebo marketing, kdy jsou vizualizovaná data formou grafů klíčová pro předpovídání trendů a včasné reagování na situaci trhu. Vizualizace si najde cestu do života všech lidí žijících v dnešní propojené společnosti. Výšková mapa s převýšením při plánování cesty, fitness náramek, systémy na monitorování glukózy, bankovní aplikace s přehledem osobních financí, jízdní řád s odjezdy MHD. Tam všude se vyskytují data ve zpracované a vizualizované podobě.

Množství informací, které je náš mozek nucen každý den zpracovávat, je neskutečné a dalo by občas říct, že je lidský mozek vnímá, zpracovává a filtruje jako stroj. Nejčastěji tedy získáváme informace z prostředků vizualizace ve formě

různých grafů (EKG pacienta), diagramů (vývoj populace), map (mapa terénu s výškovými profily) nebo například symbolů (matematické konstanty e, pí).

3.2 **Důvody, proč vizualizovat data**

Široké rozšíření a zastoupení vizualizace dat v takřka všech oblastech života je dáno zejména faktem, že lidský mozek je lepší ve vyhledávání a porozumění informací ve vizuální podobě. Takto vyobrazená data dokážeme zpracovávat výrazně rychleji než při práci s textem, tabulkami nebo čísly. Při dobré vizualizaci jsou tak skutečnosti ihned jasné. Kromě již známých skutečností lidem vizualizace často umožní zjistit i zcela nové informace a skutečnosti, o kterých předtím nevěděli, a které by jinak šly těžko vyčíst z jinak nepřehledných a neuspořádaných dat. Toto umožňuje velice rychle na tyto skutečnosti efektivně reagovat, což je v dnešní velice proměnlivé době důležité.

Asi polovina lidského mozku se zabývá zpracováváním vizuálních vstupů a velká část tohoto procesu probíhá paralelně a kontinuálně. Kupříkladu barvy budov v okolí a rychlost pohybu objektů, které nás obklopují, jsou příklady primitivních atributů, které vnímáme. Pokud je cílem vizualizace přesně zprostředkovat informace pomocí schémat, pak je nezbytné brát v úvahu percepční schopnosti člověka. Musíme vzít v potaz schopnosti uživatele porozumět informacím a také omezení lidského zrakového systému. Nedávalo by příliš smysl přetvářet proměnnou na grafický prvek, který budou mít lidé omezenou schopnost pochopit nebo kvantifikovat. [7]

Vizualizace dat v soukromé a komerční sféře je naprosto klíčová. Společnost SAS, která se věnuje vývoji softwaru pro datovou analytiku a nabízí možnost individuálních řešení pro nejrůznější oblasti podnikání, uvádí nejzásadnější důvody, proč vizualizovat data v komerční sféře. Upraveno z [8]

Vizualizace dat pomáhá předpovídat trendy na trhu a získat užitečná vodítka pro rozvoj podnikání.

Vizualizace dat umožňuje vedoucím pracovníkům na základě prognóz vycházejících z nashromážděných firemních údajů předpovídat trendy na trhu. Pochopení důležitých souvislostí pomůže získat konkurenční výhodu.

Vizualizace dat umožňuje poznat dynamiku vašeho trhu jako nikdy předtím.

Interpretace vizualizovaných dat vedením firmy má pozitivní dopad na „inteligentní“ sledování referenčního trhu. Díky tomu mohou zjistit postavení firmy na trhu v rámci daného odvětví

Vizualizace dat vám umožňuje poznat potřeby každého zákazníka a jednat podle nich.

Vizualizace je klíčová pro prodejní a marketingové aktivity. Sledování komunikačních kanálů se zákazníky postupně přináší zvýšení efektivity prodeje. Lze zkoumat data z různých zdrojů: kamenného obchodu, call centra, e-shopu nebo sociálních médií.

Vizualizace dat umožňuje přijímat správná rozhodnutí ve správný čas a podporuje rychlou výměnu informací.

Okamžitá interpretace vizualizovaných dat má za následek zefektivnění pracovních procesů a výstupů. Aplikace rozhodnutí v reálném čase přináší zefektivnění produktivity.

Dá se říct, že vizualizovaná data jsou zkrátka mnohem srozumitelnější, a pokud jsou navíc prezentována formou, která upoutá toho, komu je prezentujeme, mohou se stát vizualizační nástroje šikovnými pomocníky při vzdělávání skupin lidí, nebo pomohou při prezentaci našich názorů. Dobře vizualizovaná data upoutají diváka a získají si jeho pozornost.

3.3 Historie vizualizace dat

Techniky pro zaznamenávání informací jsou známy už z pravěku. V jeskyni Chauvet-Pontd'Arc v jižní Francii bylo nalezeno více než 250 nástěnných maleb z doby zhruba před 30 000 lety. Nejstarší systémy pro psaní využívaly obrázků a symbolů, které zastávaly funkci celých slov. Nejstarším psaným dokumentem je zřejmě kamenná deska z Kishe ze Sumeru, i když v tomto případě jsou znaky do desky vytesány. [7]

Během starověku pokročily mnohé původní civilizace s rozvojem písma, matematiky a s rozvojem jejich společností byla stále častější snaha zaznamenávat nejrůznější druhy dat, obchodní záznamy, periodické přírodní jevy, údaje o stavech armády. Situace ve středověku byla dosti obdobnou. Život bez výtvarných moderní techniky byl mnohem náročnější, den v lidském životě sestával téměř výhradně z práce vyžadující naprosto minimální práci s daty. Vizualizace dat nebyla důležitá pro obyčejné lidi. [7]

Nemůžeme však říct, že se oblast vizualizace v takzvaných „dobách temna“ vůbec nerozvíjela. Do té doby byla využívána především v kartografii, pro účely vyobrazení měst, cest a přírodních zdrojů. V 18. století se začalo využívat tematických map. „Tematická mapa (anglicky thematic map) je mapa zobrazující na podkladu základní mapy další přírodní, sociálně-ekonomické a technické objekty a jevy a jejich vztahy“. Jednoduchým příkladem může být mapa státu po volbách, kde jednotlivé části státu jsou zbarveny podle vítězných subjektů. [7][9]

Vizualizace byla ve své prvotní formě také hojně využívána při zakreslování anatomických poznatků v oblasti medicíny. Existuje spousta dochovaných knih, které sloužily pro záznam obchodních a finančních aktivit – nákup a prodej, ceny, daně a jiné. Tyto knihy mívají stránky často uspořádané do tabulkové struktury pro lepší přehled. Snad největší využití nacházela vizualizace dat ve stavitelství. Dokonale propracované a zaznamenané architektonické výkresy a postupy například pomohly Římanům vést pitnou vodu stovky kilometrů od zdroje pomocí akvaduktů s přesně vypočítaným sklonem v řádech desetin procent. Stavby z období starověku a středověku tak odolávají proudu času a působení prostředí a mnoho z nich je dodnes ve velmi dobrém stavu, mnohé i bez oprav a údržby. [7]

Do 17. století ale ani tak nebyla vizualizace příliš rozšířeným oborem. S růstem lidské populace, objevováním nových území a rozvojem vědy však nutnost vizualizovat začíná být velice naléhavá. V této části práce zmíní klíčové osobnosti, které přispěly svojí prací k rozvoji vizualizace dat, a mnoho technologií vycházejících z jejich práce je dodnes používáno na denní bázi.

William Playfair (1759–1823)

William Playfair pocházel ze Skotska. Někdy je označován za „otce grafické statistiky“. Byl povoláním inženýr a ekonom. V roce 1768 vydal svoje dílo s názvem *The Commercial and Political Atlas*, ve kterém publikoval množství nových schémat, grafů a diagramů. Je považován za vynálezce mnoha vizualizačních prvků, které dnes běžně používáme. Řadí se mezi ně čárové, sloupcové nebo výsečové diagramy. [10]

Florence Nightingale (1820–1910)

Florence Nightingale byla anglická ošetrovatelka, která během krymské války pracovala jako sestra v polní nemocnici a uvědomila si problém špatných hygienických podmínek a nevhodné stravy nemocných vedoucí k podvýživě. Kombinace mnoha faktorů vedla v vysoké úmrtnosti vojáků v nemocnici a proto začala zaznamenávat data o mortalitě a vizualizovat je. Její takzvané Rose diagramy (forma papskového diagramu) pomohly při přesvědčování vojenských představitelů, aby zajistili v lazaretech lepší podmínky. Tím zachránila mnoho vojáků, kteří by jinak zemřeli na infekce, kterým lze předejít. [10]

John Snow (1813–1858)

Když v roce 1854 v Soho v Londýně vypukla epidemie cholery, situace se na první pohled nezdála nijak vymykající se kontrole. Nákaza se ale začala šířit neuvěřitelně rychle a to navzdory opatřením, jako byla izolace nakažených, které jinak proti šíření pomáhaly. V oblast okolo ulice *The Broad Street* zemřelo ve velice krátké době na nákazu cholerou více než 600 lidí a zbytek obyvatelstva z větší části ve strachu z nákazy utekl. Snow na jednoduché mapě oblasti zanesl místa s největším počtem mrtvých. Počty mrtvých byly do mapy zaznamenány jako paralelní s ulicí na dané adrese. Díky tomuto postupu se mu podařilo najít zdroj nákazy – veřejnou studnu s pumpou na *Broad Street*. V té době nebyly hygienické standardy vůbec zavedené a speciálně to platilo v této části Londýna, která byla hustě obydlená především chudými lidmi. Proto se, s největší pravděpodobností ze splašků a odpadků běžně se vyskytujících na ulici a v okolí studny, dostala do zdroje pitné vody bakterie cholery. Lidé v té době neznali koncept patogenů, a proto

3.4 Využití rozšířené reality pro vizualizaci dat

Rozšířená realita (AR) je koncept, kdy dosadíme do skutečného okolí digitální objekty a ty pak pozorujeme na obrazovce zařízení. Během poslední doby se augmentovaná realita, v češtině zvaná rozšířená realita, stala tématem, kterému se věnují výzkumníci po celém světě. I v běžném životě rozšířenou realitu využívají miliony lidí každý den, primitivní formou AR jsou filtry fotoaparátů integrované v aplikacích sociálních sítí. Rozšířená realita si našla cestu již i do praxe, jako příklad bude uvedeno, jak se této technologii zhostila zřejmě největší česká firma ŠKODA AUTO. [11]

Technologie AR vykresluje počítačově generovaný obsah ve fyzickém okolí uživatelů. Aby se zvýšila efektivita a produktivita procesů, výzkumníci a vývojáři věnují stále větší pozornost aplikacím pro AR. Zavedením vizualizačních technologií založených na AR se překrývají výsledky inženýrské analýzy a simulace přímo na objektech reálného světa. Správná a efektivní vizualizace těchto dat pomocí aplikací pracujících s technologií AR může snížit nesprávnou interpretaci dat v prostorových aspektech. Rozvoj mobilních platforem využívajících AR technologie poskytuje inženýrům pohodlný přístup k relevantním informacím přímo na místě bez prodlevy. AR systémy již byly implementovány v biomedicíně a chirurgii. Data shromážděná magnetickou rezonancí jsou běžně vizualizována v prostředí AR. Kvůli omezené přesnosti ve zdravotnictví rozšířená realita slouží hlavně jako vzdělávací nástroj. [11]

Ve srovnání s obory jako je zdravotnictví najde AR uplatnění hlavně v civilním a městském inženýrství (Vizualizace výsledků teplotních měření a výpočet proudění tekutin – Fluid Dynamics). S pomocí AR technologií mohou stavební inženýři a městští designéři zkoumat simulované výsledky v reálném venkovním prostředí a pomocí takto získaných dat pak zlepšit design prostředí. V oblasti strojírenství a výroby je nutný okamžitý sběr dat a aktualizace parametrů reálném čase. Sběr dat probíhá nejčastěji za pomoci sítě senzorů a pracovaná data pak software zobrazí na požadovaném zařízení. [11]

V kontextu této práce by se dala uvažovat aplikace AR na pracovišti. Technici by například mohli vidět rozebranou počítačovou komponentu na monitoru, a tak provést drobné opravy bez nutnosti rozebírání celé součástky. V případě, že by se firma rozrostla a sklad komponent se přesunul do větších prostor, najde se pro AR další využití. Na obrazovce bychom viděli přesnou polohu požadovaných komponent ve skladu a zefektivnila by se tak kompletace zakázek

3.4.1 Praktické využití rozšířené reality ve ŠKODA AUTO

Spolu s kreativním studiem Brainz Immersive spustila ŠKODA AUTO testovací projekt využívající rozšířené reality. Bude sloužit prediktivní údržbě výrobních zařízení. V rámci celopodnikové digitalizace automobilka testuje moderní technologie v souvislosti se školením zaměstnanců. Nová technologie by měla přispět ke zjednodušení pravidelné kontroly výrobních zařízení a pomoci s prevencí nečekaných zastavení výrobního procesu. [12]

Pracovníci budou mít k dispozici brýle Microsoft HoloLens 2. Pomocí brýlí naskenují vybraný QR kód umístěný na stroji. Do reálného prostředí se se pak promítne seznam úkolů. Výhodou je, že interaktivní návod doplněný o fotografie a videa s pracovními postupy popisuje podrobně celou údržbou stroje. Proces údržby pak s využitím AR zabere výrazně méně času oproti práci s tištěným manuálem. [12]

Dalším příkladem použití AR v této společnosti je aplikace pro přenosná zařízení s kamerou, která po naskenování kódu kontejneru zobrazí obsah kontejneru a za pomocí AR i přesnou polohu součástí v kontejneru a zbývající volnou kapacitu. Zaměstnanci logistického úseku vidí reálnou pozici výrobků i jejich stav v jinak neprůhledném kontejneru na obrazovce zařízení. Tím se jim povedlo využít a zaplnit mezery, které v kontejnerech vznikaly. Zefektivnili tak logistiku součástí mezi výrobními závody, které se nachází například v Indii nebo Číně, a závodem v Mladé Boleslavi, kde se ze součástí kompletují hotové vozy.

4 Volba způsobů řešení

V této části se práce bude věnovat popisu a srovnání několika vybraných nástrojů. Prezentované technologie byly vybrány s ohledem na to, k jakému druhu vizualizace se používají. Také zde budou uvedeny zásady k vytvoření dobré vizualizace, nebezpečí plynoucí ze zavádějících vizualizací a budou vysvětleny rozdíly v základních pojmech. Tato část také představí vybrané diagramy, z nichž některé budou použity v praktické části.

4.1 Jak správně vizualizovat data

Pokud chceme, aby byla vizualizace opravdu přínosná a výsledná data byla zobrazena správně, musíme ji hned od začátku správně nastavit – musíme se správně ptát, abychom dostali tu správnou odpověď. Špatně formulovaná otázka může ve svém důsledku znamenat neúčinný nebo dokonce kontraproduktivní výsledek.

Pro výběr toho správného diagramu se dá použít například metoda zahrnující položení si několika základních otázek. Na webových stránkách společnosti Abra, zabývající se oblastí business intelligence a vyvíjející software, který umožňuje vytvářet množství diagramů, jsou uvedeny tři základní otázky, které bychom si měli položit při výběru vhodného vizualizačního prostředku. Převzato z [13]:

Jaké informace mají vizualizovaná data sdělit a komu?

Nejprve se musíme zaměřit na to, co se potřebujeme dozvědět. Co nás k vizualizaci vede? Kdo bude z výsledných vyobrazených dat čerpat? Každý člověk ve firmě má na starosti něco jiného, proto nemusí být stejně vizualizovaná data přínosná všem.

S jakými daty a souvislostmi pracovat a jak data připravit?

Důkladně se musíme zamyslet nad tím, jaké skutečnosti, informace, čísla či návaznosti je třeba vzít v úvahu, aby měl výsledek skutečnou vypovídající hodnotu a dalo se na něj spolehnout. Nemělo by smysl vizualizovat data jenom proto, aby pak musely být výsledky analýzy podrobovány novému zkoumání a kontrole případných chyb. Je klíčové shromažďovat pouze skutečně relevantní informace pro dané účely, zároveň by neměla být žádná relevantní informace vynechána.

Jaký vizualizační nástroj použít?

Typ grafu či barevnost tedy nejsou jen otázkou estetiky, ale mají především praktický význam. Musí být na první pohled zřejmé, na co se díváte a co je důležité. Roli přitom hrají barvy, pozice, velikost, poměry, tvary i celkové rozložení jednotlivých prvků. Některé typy grafů se hodí lépe k predikci trendů, jiné nabízejí lepší přehled o hodnotách jednotlivých kategorií. Pro jeden typ dat se může hodit více typů grafů a naopak.

4.1.1 Metriky vizualizace

Metriky vizualizace existují proto, že je potřeba nějakým způsobem měřit kvalitu informací, které nám daná vizualizace prezentuje. Základní metriky, které uvádí [14], jsou dvě: Expresivita a efektivita

Expresivita – Expressiveness – prezentuje danou informaci, klíčové je vybrat pro vizualizaci taková data, která neobsahují chybnou informaci. Pomocí expresivity se měří koncentrace informace v dané vizualizaci.

Efektivita – Effectiveness – vizualizace je efektivní, když je možné ji interpretovat přesně a rychle a jestliže je možné ji zobrazit s využitím adekvátních zdrojů (výpočetní výkon, čas...)

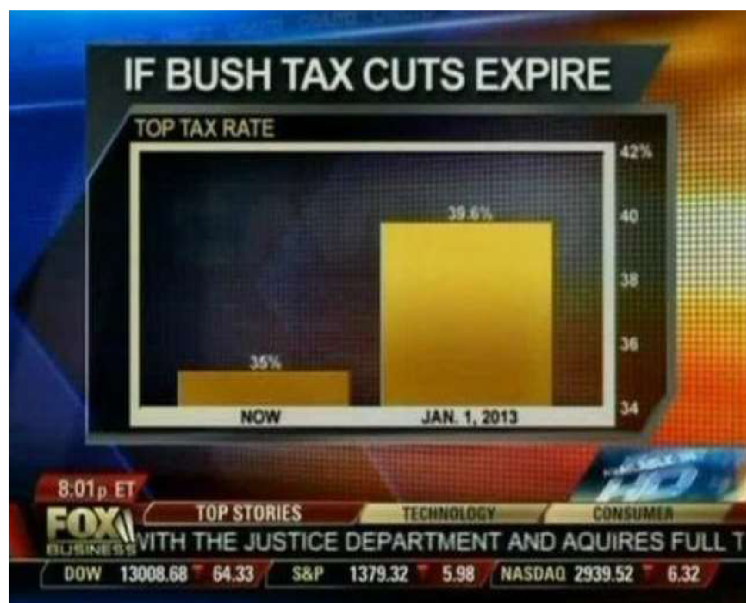
4.2 Nebezpečí zavádějící vizualizace dat

Zavádějící statistiky a z nich vycházející vizualizace mají někdy záměrně ovlivnit veřejné mínění a někdy je to jen případ, kdy lidé nerozumí datům a špatně si je interpretují. Nebezpečí zavádějících statistik je bezprostřední tedy zejména při ovlivňování veřejného mínění, volebních výsledků a podobně. Zejména když se vzhledem k propojené společnosti mohou takové zavádějící informace dostat, a velmi často se také dostávají, k lidem náchylným k tomu jim uvěřit. Dále v této kapitole jsou uvedeny příklady zavádějících statistik a grafů publikovaných v médiích a politiky. Toto jsou nejčastější případy chyb v zavádějících grafech podle [15]:

Jedno z měřítek je příliš velké nebo příliš malé, přeskakuje čísla nebo nezačíná od nuly

Graf není správně označen

Důležitá data jsou vynechána.

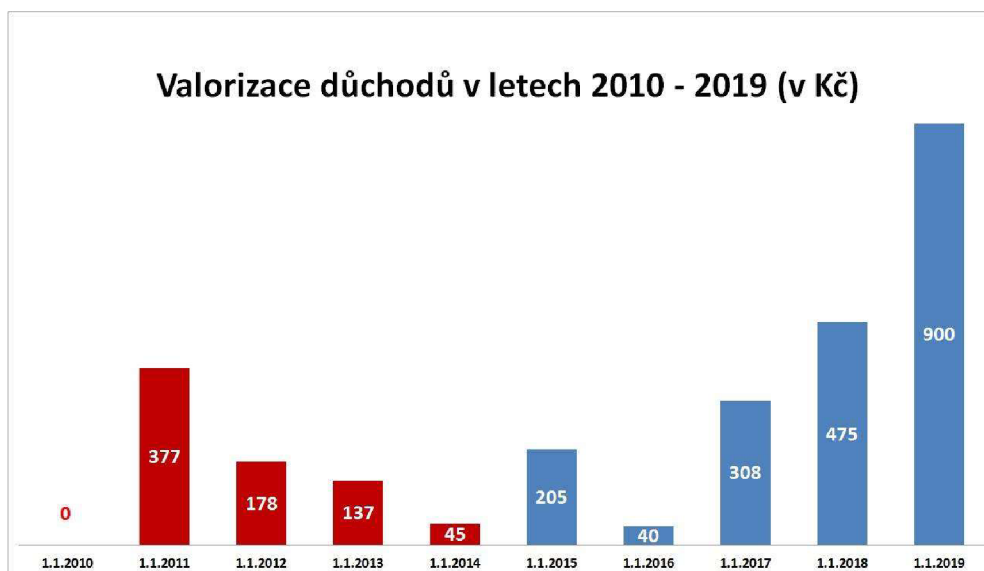


Obrázek 3 – zavádějící statistika Fox News [18]

Takto na Fox News prezentovali, co by se stalo, kdyby vypršely v té době platné daňové tarify. Osa diagramu nezačíná od nuly. Proto to na první pohled vypadá jako velký nárůst, ale ve skutečnosti je záměrně uvedená zavádějící škála. Nárůst byl ve skutečnosti 4,6 procent, ale na první pohled se může zdát, jako by se daňové tarify několikanásobně zvedly. [15]

Další příklad můžeme uvést přímo z prostředí české politické scény. Data a grafy prezentovány minulou vládou byly často v absolutních hodnotách, (milionech korun a podobně), a ne relativních, přepočtených jako procenta vůči HDP nebo jinému ukazateli. Absolutní hodnoty, co se týče peněz, podléhají inflaci a jiným faktorům. Na uvedeném diagramu je vidět průměrné zvyšování valorizace důchodů. První problém je ten, že není zohledněn faktor inflace a s tím spojený souběžný růst mezd a zvyšování cen zboží. (v roce 2019 výrazně vzrostly oproti minulým rokům ceny za energii a potraviny) Další problém je uvedení průměrné hodnoty růstu důchodů, která je ovlivněna nízkým počtem vysokopříjmových jedinců a na kterou často většina lidí nedosáhne.

V jiném případě byl publikován graf, který ukazoval zvýšení průměrného platu policistů v roce na 44 476 Kč, což bylo v rozporu s 40 820 Kč uvedenými na webu policie. [17]



Obrázek 4 – zavádějící schéma české vlády [16]

4.3 Schémata, grafy a diagramy

4.3.1 Schéma

Schémata (anglicky charts) jsou prostředky používané k vizuálnímu vyobrazení informací. Schémata mají mnoho podob. Takzvaná nástěnná schémata (wallcharts) mohou sloužit ke vzdělávacím a informativním účelům, jejich součástí mohou být třeba i náčrty a plány budov. Ganttův diagram se používá k rozdělení práce na úkolech, koláčový graf použijeme k jasnému sdělení rozdělení hodnot v souboru. Pojmy chart, diagram a graf se jak v odborné literatuře, tak v praxi často zaměňují a používají jako synonyma. [18][19]

4.3.2 Grafy a diagramy

Tato práce definuje pojem graf z pohledu teorie grafů a technologie využívané při vizualizaci budou označovány jako diagramy. Většina lidí si s největší pravděpodobností pod pojmem graf představí náčrt klasického dvourozměrného prostoru s vyznačenými osami x a y, který pomocí spojení ukazuje proměnu hodnot v čase. Ale rovněž je graf i matematický koncept. Takovéto grafy jsou ucelené

soubory sestávající z bodů. Tyto body nazýváme nejčastěji vrcholy nebo uzly. Vrcholy a uzly bývají propojené. K tomu se používá takzvaných hran (edges) a odkazů (links). Matematicky by se tedy dal velice jednoduše graf vyjádřit takto: **Graf G = (Vrcholy V, Hrany H)** Grafy mohou být orientované a neorientované. Jak hrany, tak vrcholy mohou mít přidělené číselné skóre reprezentující reálné hodnoty. V případě orientovaných grafů se zakresluje v grafu šipka. Grafy mohou být souvislé a nesouvislé. Hovoříme-li o souvislých grafech, znamená to, že z každého vrcholu A existuje spojení s vrcholem B (C, D...). [18][19]

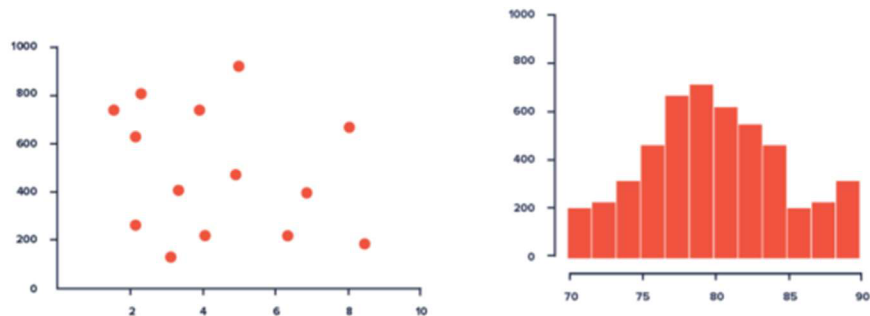
Pokud chceme, aby druhá osoba pochopila informace, které se jí snažíme předat, musíme prezentovat data stručně a jasně. K tomu nám poslouží diagramy. Diagramů je nespočet druhů a jejich individuální použití je vždy závislé na mnoha faktorech, jako je druh dat (čísla, řetězce), množství hodnot a odchylky mezi nimi.

Definice a schémata k prezentaci jsou převzaty ze stránek Data Viz Project. Uvedené a popsané níže jsou ty nejčastěji používané druhy grafů a diagramů. Pro větší přehlednost jsou grafy rozděleny do čtyř skupin tak, jak je uvádí Data Viz Project [20]:

4.3.2.1 Diagramy založené na kvantitativních datech

Korelační diagram – v angličtině se používá termín scatterplot. Je to matematický diagram využívající kartézské souřadnice k zobrazení hodnot dvou proměnných. Data jsou zobrazena jako soubor bodů, každý bod má polohu určenou hodnotou na vodorovné ose a hodnotou druhé proměnné na svislé ose. [20]

Histogram – diagram seskupující číselná data do intervalů stejné šířky, velikost každé skupiny je zobrazena pomocí sloupců. Intervaly jsou určeny předem podle druhu vizualizovaných dat. Může jít o věkové skupiny, teplotní rozsahy a podobně. Histogram je tedy sloupcovým grafem, ale oproti němu u něj platí právě zásada, že data na ose x jsou seskupeny do kategorií podle určitých číselných intervalů. Takzvaný barevný histogram najde využití v počítačové grafice a digitální fotografii, kde zobrazuje intervaly frekvenčního rozsahu viditelného spektra světla ve vztahu k výskytu různých barev. Sloupce v tomto případě vyjadřují počet pixelů snímku spadajících do daného intervalu spektra. [20]



Obrázek 5 – Korelační diagram a Histogram [20]

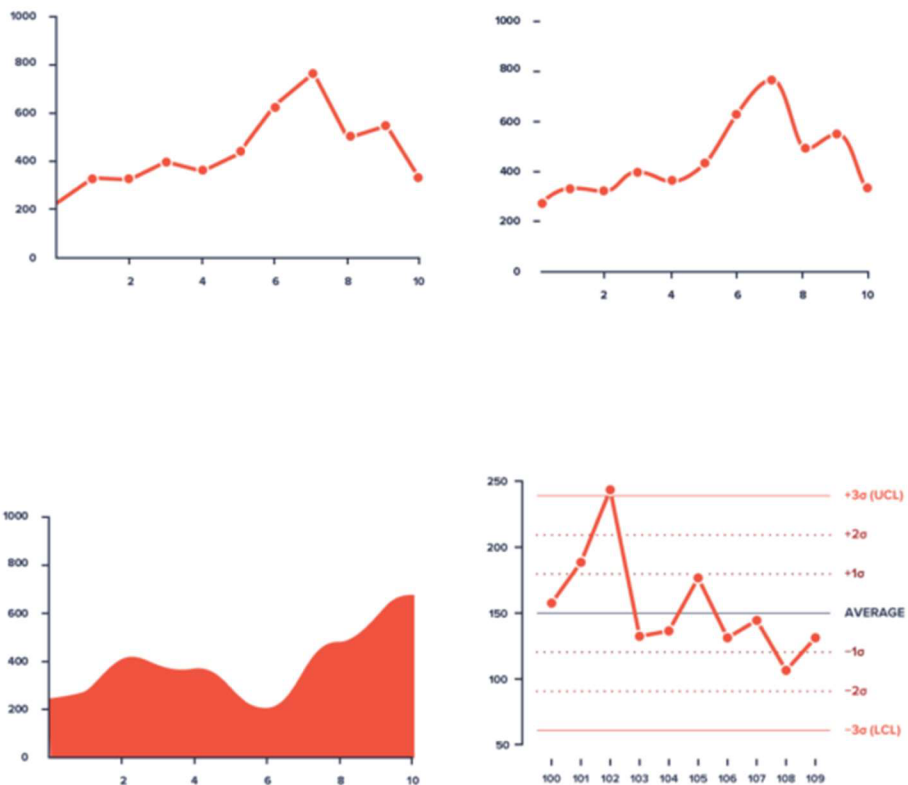
Regulační diagram – jinými slovy také kontrolní diagram. nástroj používaný při řízení procesu. Je to spojnicový diagram proměnné z procesu, který je doplněn o kontrolní vodící linky, které určují, jak moc se proměnná během procesu odchyluje od výchozího (průměrného) stavu. Při vytváření kontrolního diagramu je nejprve vybráno referenční časové období a z hodnot za toto období se vypočítá průměr. Ten se zanesse jako vodorovná čára uprostřed grafu. Nad i pod průměrem se vypočítají tři směrodatné odchylky, horní kontrolní limit (UCL) a dolní kontrolní limit (LCL). Při čtení grafu se zjišťují významné odchylky. Například hodnotám ležícím mimo intervaly limitů je nutno věnovat zvýšenou pozornost, stejně tak dlouhým řadám po sobě jdoucích hodnot nad nebo pod průměrem. Regulačního diagramu je hojně využíváno při kontrole výstupu průmyslových výrobních prostředků. [20]

Liniový graf – jinak také spojnicový, je obdobou bodového grafu, ale rozdíl je v tom, že body zanesené na grafu jsou seřazeny a spojeny jednoduchými linkami. Liniový graf vizualizuje, jak se mění určitá data v časových intervalech. Využívají ho srovnávače cen na internetu, často se objevují body a linky vyobrazující průměrnou cenu a minimální cenu produktu. [20]

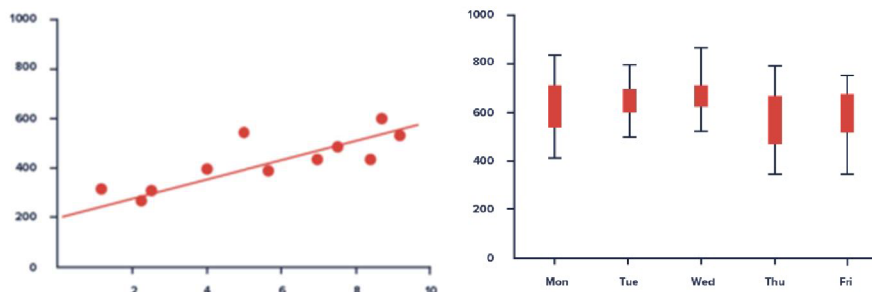
Spline chart – je to obdoba, a dalo by se říct vyhlazená verze, klasického liniového grafu. Na rozdíl od něj jsou ve spline diagramu vrcholy protnuty křivkou a díky tomu je možné přesněji analyzovat vývoj trendu z diagramu. Stejně jako liniový diagram zobrazuje proměnu dat během stejně dlouhých intervalů. [20]

Plošný graf – modifikace liniového (spline) diagramu. Oblast pod křivkou je vyplněna barvou. Ve statistice se užívá při vizualizaci výpočtu oblasti pod křivkou, může zobrazovat celkové objemy výroby a podobně. [20]

Svícový graf – vyobrazuje například pohyby kurzů měn u různých finančních institucí za zvolené časové období. Svícový graf obsahuje horní a spodní hodnotu, reprezentovanou vodorovnou linkou. Vizuálně je velice podobný boxplot diagramu. [20]



Obrázek 6 –Liniový, Spline a Plošný graf, Regulační diagram [20]



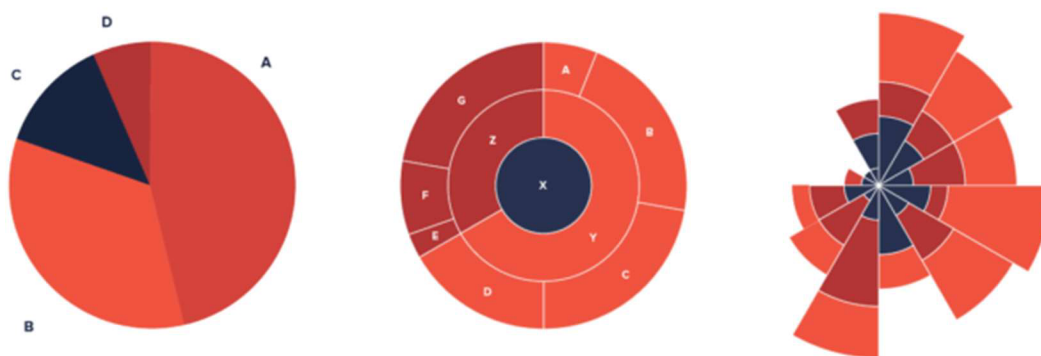
Obrázek 7 – Trendline a Svícový graf [20]

Trendline – doslova přeloženo trendová čára. Modifikovaná verze scatterplotu. Tento diagram obsahuje navíc linku, která zobrazuje obecný průběh nebo trend zkoumané situace. Linka je jakýmsi „zprůměrováním“ číselných hodnot vrcholů grafu. Používá se zejména ve financích. [20]

Kruhový diagram – kruhový, někdy také koláčový, graf je rozdělený na sektory vyobrazující kvantitativní (nejčastěji procentuální) zastoupení jednotlivých prvků v souboru. Jméno získal podle toho, že jednotlivé rozdělené segmenty vypadají jako nakrájený koláč. [20]

Víceúrovňový kruhový diagram – umožňuje podrobněji rozdělit dané hodnoty. Praktický příklad: procentuální zastoupení dvou hlavních skupin (pohlaví muž/žena) je vyobrazeno ve vnitřní části diagramu a vnější část je rozdělena na další podskupiny (věk, dosažené vzdělání). [20]

Polar Area Graf – podobný kruhovému grafu. Používá se při prezentaci cyklických dat jako porodnost a mortalita nebo změna teploty během roku. Každý segment v něm může mít různou délku od středu. Tyto segmenty se dají přirovnat ke sloupcům u sloupcového grafu. Rozměry každého segmentu udávají celkový poměr buď v daném „sloupci“, nebo v rámci celého diagramu. Poprvé ho využila Florence Nightingale při prezentaci mortality raněných vojáků a získala díky tomu prostředky pro zlepšení podmínek v lazaretech. [20]



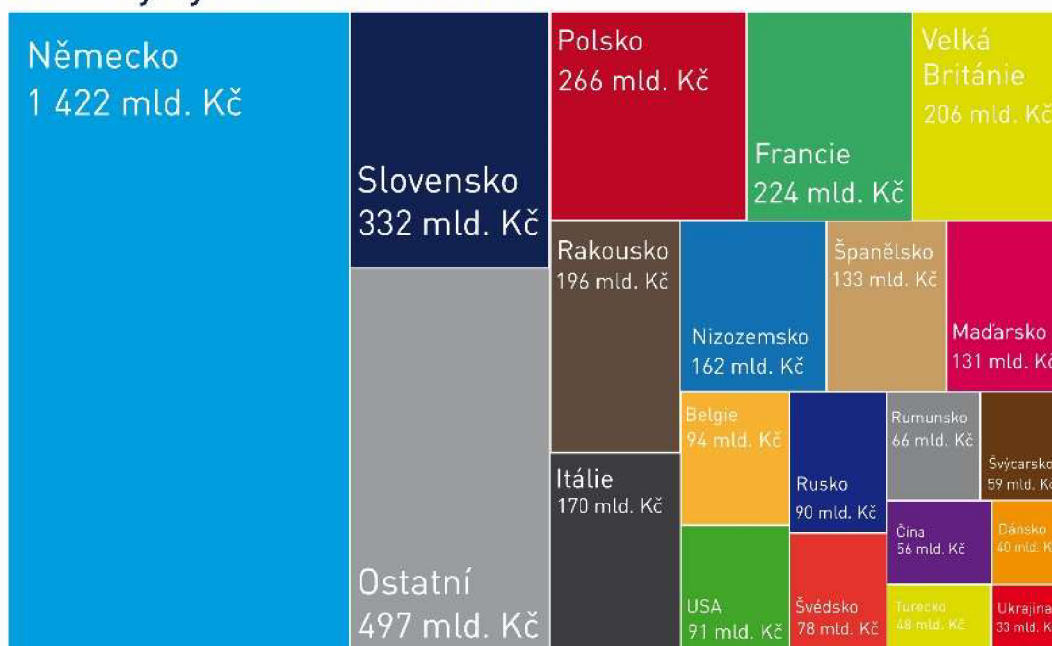
Obrázek 8 – Kruhový diagram, Víceúrovňový kruhový diagram a Polar Area Graf [20]

Stromová mapa – Zobrazuje data hierarchicky ve formě obdélníkových segmentů. Části stromové mapy jsou obdélníky děleny dále a vzniká tak podrobnější pohled na data. Velikost každého segmentu je proporční k určitému množství dat. Nejčastěji je tento diagram používán v ekonomice, kde lze pomocí něho přehledně vyobrazit například složení exportu daného státu. Na následujícím příkladě publikovaným Svazem průmyslu a dopravy ČR je vidět, jak je důležitá vhodná volba kontrastních barev v grafech. Některé popisky jsou nečitelné. [20]

KAM VYVÁŽELY ČESKÉ FIRMY V ROCE 2018



Celkový vývoz: 4 391 mld. Kč

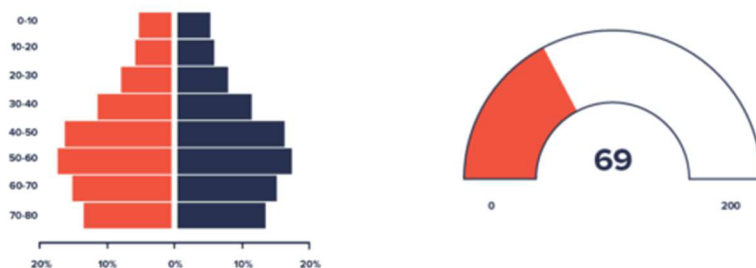


Zdroj dat: ČSÚ

Obrázek 9 – Stromová mapa exportu firem v ČR [21]

Populační pyramida – jinak zvaná věková pyramida, je grafickou ilustrací distribuce určených věkových skupin ve zkoumané populaci. Pokud populace roste, tvoří tvar pyramidy. [20]

Solid Gauge Chart – vyobrazuje škálu hodnot na zaoblené stupnici. Často se s obdobným schématem můžeme setkat u tlakoměrů, kompresorů a podobné techniky kdy nám ručička na ciferníku ukáže tlak. Může být přítomno barevné značení optimálního tlaku. [20]



Obrázek 10 – Populační pyramida a Solid Gauge Chart [20]

4.3.2.2 Diagramy založené na nominálních datech

Word Cloud – jinými slovy také tag cloud. Vizualizace textových dat, která se obvykle používá k zobrazení tagů na webových stránkách, k vizualizaci volného textu nebo jako vizuální podpora různých kampaní. Důležitost každého tagu je znázorněna velikostí písma nebo barvou. Tento formát je užitečný pro rychlé vnímání nejvýznamnějších prvků a hesel. Při pozorování snadno vizuálními vlastnostmi tagu určíme jeho důležitost. [20]

Cyklický diagram – použití najde v průmyslových výrobních procesech, jednoduše popisuje události tak, jak jdou za sebou a opakují se v rámci cyklu. Celkový průběh procesů a události uvedené v cyklickém diagramu nemusí mít vyznačený přesně konec ani začátek. [20]

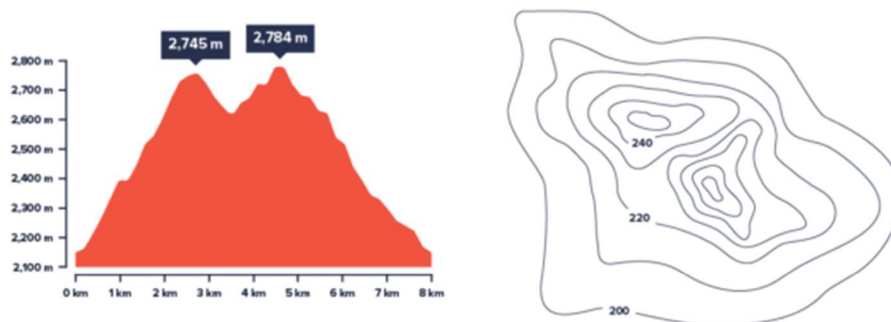


Obrázek 11 – Word Cloud a Cyklický diagram [20]

4.3.2.3 Geoprostorové diagramy

Profilová mapa – pohled na geografická data, který se zaměřuje na nadmořskou výšku. Příčný řez terénem zobrazující nadmořskou výšku opisující cestu. Hodnoty nadmořské výšky jsou zaneseny podél svislé osy, zatímco na vodorovné ose je vyznačena vzdálenost podél cesty. Je jedno, zda se jedná o leteckou čáru protínající prostor, nebo o složitější cestu vedenou podél terénu. Kvůli zdůraznění výškového profilu terénu bývá vertikální měřítko oproti horizontálnímu mnohokrát zveličené. Mapu profilu lze vytvořit opisem čáry řezu na topografické mapě a zanášením hodnot nadmořské výšky, když čára protne vrstevnice. [20]

Topografická mapa – Topografická mapa je podrobné a přesné grafické znázornění nadmořské výšky terénu na zemi. Topografická mapa sestává z vrstevnic. Vrstevnice je liniový segment, který je spojený a tvoří uzavřený okruh. Jednotlivé vrstevnice se ale vzájemně neprotínají. Každá vrstevnice představuje nadmořskou výšku. Různou vzdáleností mezi sousedícími vrstevnicemi se dá určit převýšení a náročnost terénu. [20]



Obrázek 12 – Profilová mapa a topografická mapa [20]

Kartogram – Kartogramy jsou primárně používány k vizualizaci dat týkajících se zemí, regionů nebo států, například výsledky ve volbách, počet obyvatel nebo výše příjmů. Kartogramy mohou deformovat tvar geografických oblastí, takže oblast lépe představuje datovou proměnnou. Běžným příkladem je překreslení každé země na světě na jinou velikost v závislosti na HDP dané země. Ekonomicky výkonnější země se pak jeví na mapě mnohokrát větší, než ve skutečnosti jsou, a naopak rozlohou velké, méně ekonomicky vyspělé země jsou deformovány a zmenšeny. Také může být původní mapa úplně předělána formou jiných vizualizačních nástrojů, například dílčích geometrických obrazců, jejichž velikost bude proporčně vyobrazovat reálné hodnoty. Níže je vidět kartogram vyobrazující zkreslenou velikost států ve světě podle populace. [20]



Obrázek 13 – Kartogram světové populace [20]

4.3.2.4 Tabulkové diagramy

SWOT Analýza – někdy také SWOT matice. Zkratka je složená z počátečních písmen slov Strengths – Weaknesses – Opportunities – Threats. Slouží při manažerských rozhodovacích procesech jako jsou plánování a zpětná kontrola. Je to strukturovaná metoda používaná k hodnocení silných a slabých stránek, příležitostí a hrozeb. Ve SWOT analýze se také často uvádějí příznivé a negativní vlivy, které mohou být externího nebo interního původu. Běžně se používá při uspořádání podnikatelského plánu a/nebo při zakládání společností, aby byla vhodně odprezentován potenciál projektu uspět na trhu. [20]

Teplotní mapa – typ vizualizace dat, kde jednotlivé hodnoty obsažené v matici (tabulce) jsou rozděleny prostřednictvím variací ve zbarvení. Termín „teplotní mapa“ byl zaveden až v roce 1991 softwarovým designérem Cormacem Kinneyem, když popisoval 2D zobrazení informací o finančním trhu v reálném čase. Podobné vizualizace ale existují již více než století. Teplotní mapy jsou užitečné pro vizualizaci rozdílů mezi více proměnnými a pro vyhledávání vzorců a korelací. [20]



Obrázek 14 – SWOT Analýza a Teplotní mapa [20]

4.3.3 Kontingenční tabulka

Ve statistice je kontingenční tabulka také známá jako křížová tabulka. Je to matice, která zobrazuje (v některých případech taktéž vícerozměrné) rozdělení četností proměnných. Kontingenční tabulky jsou hojně používány v průzkumech trhů, business intelligence, technologických odvětvích a inženýrství a vědeckém výzkumu. Dávají nám na první pohled základní obrázek o vzájemném vztahu mezi dvěma proměnnými. Kontingenční tabulky jsou použity jako úložiště dat při tvorbě diagramů, ale již na první pohled mohou pomoci najít vzájemné vztahy, trendy a interakce v souboru dat. [22]

Příklady kontingenčních tabulek jsou uvedeny v praktické části. Původní tabulky s daty jsou často zpracovány na dílčí kontingenční tabulky, pomocí kterých se vytvářejí přehlednější vizualizace. Například v Excelu jsou kontingenční tabulky základem pro široké spektrum grafů a diagramů. Aplikace Google Data Studio

umožňuje v kontingenčních tabulkách vybírat a filtrovat pouze data splňující požadovaná kritéria a tak zobrazit pouze požadované skupiny dat a jejich počet.

5 Technologie vizualizace dat

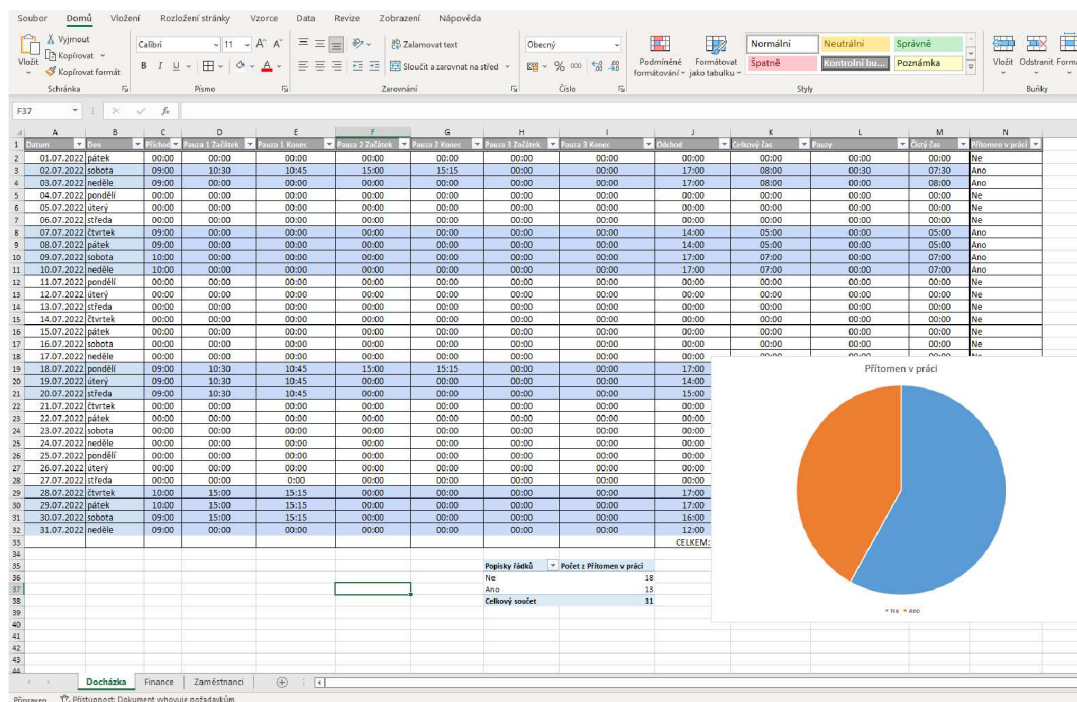
5.1 *Microsoft Excel*

Excel pracuje na principu tabulkového sešitu, kam zapisujeme data. S daty se provádí operace za pomoci definovaných funkcí programu, jsou mezi nimi matematické funkce, například SUMA nebo PRŮMĚR, a také obecné funkce programu – hledání a náhrada dat, filtrování výsledků podle zadaných kritérií. Nabízí možnost převedení určité skupiny dat na kontingenční tabulku a pomocí této provádět s daty další operace nebo je vizualizovat.

Excel je nejrozšířenějším kancelářským tabulkovým procesorem. Je součástí základního balíčku aplikací kancelářské sady MS Office. Využívá se jak v domácnostech, tak je hojně rozšířen pro kancelářskou činnost v podnicích, správu financí, produktů a podobně. S tím neodbytně souvisí potřeba vizualizovat data zpracovávaná Excelem. Excel má tuto funkcionalitu integrovanou, a tak firma může využívat vizualizačních nástrojů (grafů, diagramů, vizuální úpravy dat) bez potřeby zavádění dalšího specializovaného nástroje.

Mezi vizualizační prvky, které Excel umožňuje za pomoci integrovaného nástroje vytvořit z dat v tabulce, patří rozličné sloupcové, kruhové nebo liniové diagramy. Data přímo v tabulce se také dají pro větší přehlednost za pomoci nastavení buněk upravovat tak, aby působila lepším vizuálním dojmem – barvy textu, velikosti, zvýraznění buněk a ohraničení.

Data se do Excelu mohou naimportovat z databáze, nebo z různých souborů. Při importu ze souborů se dají nastavit znaky oddělující data v tabulce. Excel umožňuje pokročilejším uživatelům definovat makra a tím zefektivňovat práci s daty v rámci programu.



Obrázek 15 – Pracovní plocha aplikace MS Excel 2016

Předchůdce aplikace Excel se nazýval Multiplan. Jednalo se o aplikaci pro MS-DOS. První verze Excelu vyšla pro Macintosh v roce 1985 a přinesla mnohé revoluční prvky. Jako první využíval rozbalovacích oken s nabídkami a měl integrovanou podporu myši, takže práce s ním byla oproti jiným programům jednodušší. Také nabízel možnost data zapisovat v 256 fontech. Kvůli těmto výhodám se rozhodlo mnoho uživatelů z jiných systémů přejít právě na Macintosh. Když byl později uveden jako jedna z hlavních aplikací nového systému MS Windows, opět mnohé uživatele přesvědčil o změně systému. [23]

V roce 1993 Microsoft přidal podporu Visual Basic (VBA) a tím umožnil vytváření nejrůznějších druhů maker. To mělo za následek i šíření takzvaných makrovirů – škodlivých kódů zabalených souborem s Excelovou tabulkou. Ve stejném roce také vyšla první verze kancelářského balíku MS Office. Od té doby má Excel dominantní postavení na trhu. Současná verze aplikace je dostupná jak pro desktopové operační systémy, tak i pro mobilní zařízení. Existuje také webová aplikace. Velká popularita Excelu způsobila, že se mnoho vývojářů snažilo vytvořit aplikaci se stejnými funkcemi. Dnes tedy Existují i freeware klony Excelu, například Apache OpenOffice, LibreOffice nebo Google Spreadsheets. [23]

5.2 **Microsoft Power BI**

Microsoft Power BI je hlavní nástroj, který bude využitý při zpracování dat v praktické části. Následující text vychází z manuálu aplikace dostupného na stránkách výrobce. Microsoft Power BI je interaktivní software sloužící k vizualizaci dat se zaměřením na Business Intelligence.

Power BI je soubor softwarových služeb a aplikací, které pomáhají uživateli přetvořit zdánlivě původně nesouvisějící zdroje dat na ucelené a přehledné vizuální prostředky, s jejichž pomocí je prezentace výsledků pohlcující a interaktivní. Zdrojem dat mohou být Excelová tabulka, textové soubory, databáze a cloudové datové sklady. Power BI umožňuje snadno se připojit ke zdrojům dat, vizualizovat je a objevovat ve zpracovaných datech důležité souvislosti. Uživatelé mají mnoho možností, jak sdílet výsledky svojí práce. [24]

Tuto aplikaci původně vytvořili Thierry D'Hers a Amir Netz ze společnosti Microsoft. Původní označení aplikace znělo Project Crescent. Ten byl původně k dispozici pro veřejné stažení 11. července 2011 v balíčku SQL Serveru Denali. Později byla ale aplikace přejmenována na Power BI, podle výrazu Business Intelligence. Představena byla oficiálně společností Microsoft v září 2013 jako Power BI pro Office 365. Postupem času Microsoft přidával do aplikace mnoho funkcí, jako jsou lepší datová konektivita. Nástroj Power BI byl poprvé uvolněn pro širokou veřejnost 24. července 2015. [25]

Power BI se skládá z několika dílčích částí, které spolu všechny navzájem spolupracují. Toto jsou ty hlavní [24]:

Desktopová aplikace pro Windows s názvem **Power BI Desktop**.

Online služba SaaS (Software as a Service) s názvem **Power BI Service**.

Aplikace **Power BI pro mobilní zařízení**.

Power BI Desktop

Jedná se o bezplatnou aplikaci, která se nainstaluje do počítače a poté umožní připojit se k datům, transformovat a vizualizovat je. Power BI Desktop se připojí k více různým zdrojům dat a kombinuje je do datového modelu. Tento datový model

pak umožňuje vytvářet rozličné vizualizace, které se pak dají snadno přímo pomocí funkcí aplikace sdílet jako sestavy s ostatními lidmi v organizaci. Power BI Desktop je často využíván datovými analytiky a profesionály v oblasti business intelligence. Aplikace se aktualizuje každý měsíc a zahrnuje opravy a vylepšení řízené zpětnou vazbu od zákazníků a nejnovější trendy. Nejnovější verzi Power BI Desktopu lze získat z Windows Store, nebo jako jeden instalační soubor obsahující všechny podporované jazykové mutace. [24]

Následuje popis základních prvků uživatelského rozhraní aplikace a první kroky potřebné k zahájení práce s Power BI. Na levé straně pracovní plochy jsou dostupná tři zobrazení:

Report: V tomto zobrazení se vytváří sestavy a vizualizace.

Data: V tomto zobrazení jsou vidět tabulky a další data použitá v datovém modelu. Je zde možnost transformovat data pro co nejlepší využití při další práci.

Model: V tomto zobrazení slouží ke správě vztahů mezi tabulkami v datovém modelu.

Existuje mnoho různých zdrojů dat, ke kterým se lze připojit, poté je možnost čistit a transformovat data pomocí integrovaného editoru Power Query – ten je i součástí MS Excel. Změny v datech, jaké lze provádět, jsou například změna typu dat, odebrání sloupců a řádků tabulky nebo kombinování dat z více zdrojů. Když je datový model hotový, dají se v zobrazení sestavy a vytvářet četné vizualizace. Kolekce vizuálních reprezentací dat v souboru Power BI Desktopu se nazývá sestava. Sestava může mít jednu nebo více stránek, stejně jako sešit v Excelu může mít jeden nebo více listů. [24]

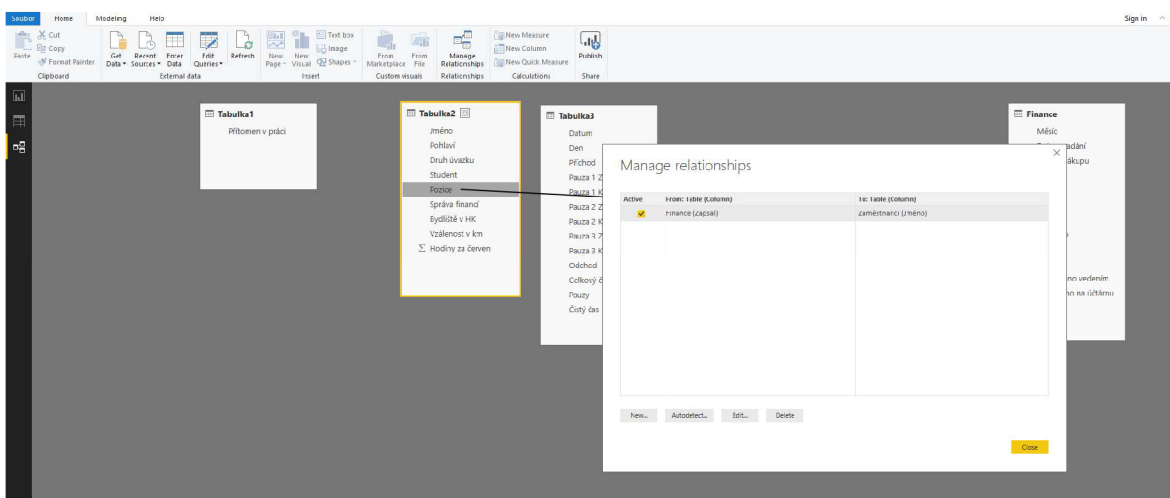
S Power BI Desktopem lze vytvářet složité a vizuálně výřečné sestavy, které pak pomocí aplikace přímo sdílíme s ostatními ve vaší organizaci. Sestavy se dají zpřístupnit komukoli s licencí Power BI. Aplikace dá na výběr, kde bude sestava sdílena, například váš pracovní tým, celá organizace nebo udělí individuální přístup konkrétním uživatelům. Také lze převést sestavy na jiné umístění v online službě Power BI Service. Pro získání Power BI Desktop, jde využít jednoho ze dvou způsobů [24]:

Stáhnout jako speciální typ aplikace z obchodu Microsoft Store.

při získání Power BI Desktopu z obchodu Microsoft Store výrobce uvádí několik výhod: Automatické aktualizace, menší objem stahovaných dat, oprávnění správce systému nejsou vyžadována, zavádění na více systémech, automatická detekce jazyka [24]

Stáhnout spustitelný soubor a nainstalovat do počítače.

Existuje také možnost stažení spustitelného instalačního souboru Power BI Desktopu z Centra pro stahování. Na výběr ke stažení je 32bitový nebo 64bitový instalační soubor. Power BI Desktop se dodává jako jeden instalační balíček ve formátu .exe, který obsahuje všechny podporované jazyky. Tento způsob distribuce je ze stran aktualizací a modifikace aplikace mnohem jednodušší a pohodlnější zejména pro administrátory. K přizpůsobení parametrů instalace lze také využít příkazového řádku. Aplikace vyžaduje se alespoň 1440x900 nebo 1600x900 a Windows 8.1 nebo novější. Nižší rozlišení nejsou podporována, protože některé ovládací prvky se zobrazují mimo tato rozlišení. [24]



**Obrázek 16 – Pracovní plocha aplikace Microsoft Power BI Desktop
Power BI Online Service**

Power BI se skládá z více softwarových služeb a aplikací propojených tak, aby umožnily vytvářet, sdílet a využívat obchodní statistiky způsobem, který uživatelům poslouží co nejefektivněji. Služba Microsoft Power BI Service

(dostupná na adrese app.powerbi.com) bývá také někdy označovaná jako Power BI online. Jde o takzvaný SaaS (Software as a Service). [24]

Prvky nazývané řídicí panely (dashboards) se zobrazují jako dlaždice, které zobrazují data a vizualizace z vybraných sestav. Řídicí panely a sestavy se připojují k datovým sadám a shromažďují všechna relevantní data na jednom místě. [24]

V typickém pracovním postupu Power BI se začne vytvořením sestavy v Power BI Desktopu a poté se publikuje do služby Power BI. Po vytvoření sestav a řídicích panelů je lze sdílet, aby je koncoví uživatelé ve službě Power BI Service a na mobilních zařízeních mohli prohlížet a pracovat s nimi. Možnost určit, jak a s kým sdělíte svou práci, je jednou z nejdůležitějších funkcí služby Power BI. Jakožto součást pracovního týmu, mohou uživatelé se svými kolegy spolupracovat na sestavách a řídicích panelech a poté je seskupovat a distribuovat jako aplikace. Také je možné sdílet samotná data a datové modely, takže je ostatní uživatelé Power BI mohou použít jako základ pro své vlastní sestavy. [24]

5.3 Google Data Studio

Hlavním účelem aplikace Google Data Studio je pomáhat uživatelům vytvářet dynamické, vizuálně působivé sestavy a řídicí panely propojením externích zdrojů dat do snadno přehledné platformy. Testovací verze byla spuštěna v květnu 2016 jako součást placené sady Analytics360 od Googlu. V srpnu 2016 Google zpřístupnil program zdarma široké veřejnosti. Následující popis aplikace vychází z publikovaného článku Product Review Google Data Studio od Genifer Snipes [26]

Základní funkcí aplikace Google Data Studio je vizuální interpretace komplexních dat shromažďovaných ze sociálních médií pomocí nástrojů webové analýzy jako jsou Google Ads a YouTube Analytics. Nicméně jeho podpora databázových nástrojů jako MySQL a jednoduchých zdrojů dat jako Google Sheets umožňují využití výzkumnými pracovníky a analytiky v komerční sféře. Ti mohou interpretovat svá vlastní data ve stejně atraktivním a uživatelsky přívětivém formátu. [26]

Schopnost Google Data Studio vytvářet a vizuálně reprezentovat klíčové metriky je velmi užitečná pro obchodní pracovníky a výzkumníky v oblasti komunikačních technologií. Využití této technologie je vhodné pro toho, kdo pracuje s daty shromažďovanými nástrojem Google Analytics. Mezi ty, kteří tyto služby využívají, patří výzkumníci z Meziuniverzitního Konsorcia pro Politické a Sociální Výzkumy (ICPSR), Organizace spojených národů (OSN) nebo United States Census Bureau. Nesporná výhoda Google Data Studio je ta, že je k dispozici zdarma pro každého, kdo má účet Google. Je tak potenciálním řešením pro řadu výzkumníků, kteří plnohodnotné programy pro vizualizaci dat nevyužijí, protože jsou buď nedostupné z důvodů vysokých pořizovacích a provozních nákladů nebo nevhodné k využití kvůli přehnané složitosti. [26]

Uživatelské rozhraní aplikace je rozpoznatelné a pochopitelné pro každého, kdo pracuje s jinými kancelářskými aplikacemi od Googlu nebo jiným podobným softwarem. Google Data Studio spoléhá na obvyklý mix vizualizačních nástrojů stejně jako ostatní vizualizační aplikace, ale rozšiřuje je o funkce jako možnost integrovat více zdrojů dat do jedné sestavy a umožňuje dynamické aktualizace výsledků. [26]

Další z působivých funkcí tohoto nástroje je umožnění uživatelům interagovat s nálezy v reálném čase pomocí filtrování výsledků na základě konkrétní proměnné. Funkce pro sdílení a spolupráci zase umožňují sdílet výsledky práce veřejně či soukromě a popřípadě nastavit přístupová pravidla, která dávají dalším spolupracovníkům možnosti prohlížet a upravovat jejich práci dle libosti. Vizuální sestavy v Google Data Studiu existují nezávisle na jejich podkladových datech, takže když vlastník sestavy uděluje přístup k úpravám nebo prohlížení dalším uživatelům může nastavit, aby tento člověk viděl pouze výsledné datové vizualizace, ale zároveň aby neměl přístup k původním zdrojovým datům. [26]

Stejně jako ostatní produkty od Googlu, je i Google Data Studio schopné zpracovávat data ve 37 jazycích a podporuje 59 světových měn. Při volbě použití Google Data Studio by měli uživatelé zvážit také typ dat, která budou zpracovávat. Pro anonymizovaná data je Google Data Studio v pořádku. Mnoho cloudových služeb, včetně té od Googlu, je ale považováno za ne úplně bezpečné. Jejich servery

nemusí splňovat požadavky IRB (Institutional Review Board – Institucionální kontrolní komise) na ochranu osobních údajů. A tak mohou být námi svěřená data v ohrožení. Nehledě na to, že celá aplikace pracuje výhradně přes internet, a tedy ani provozovatel služby ani uživatel nemohou mít zaručenou stoprocentní bezpečnost. Při práci s citlivými daty bychom měli tedy zvážit, zda je vhodné pro zpracování dat volit tuto aplikaci. [26]

Přestože nejsilnější stránkou Google Data Studio je práce s datovými podklady vytvořenými a spravovanými společnostmi Google (Google Ads, YouTube Analytics a podobně) má i potenciál využití při prezentaci vlastních dat výzkumníky. Snadno srozumitelné rozhraní a integrace s ostatními produkty Googlu jsou dalšími přednostmi. A i když Google Data Studio nabízí méně funkcí a je méně výkonné než některé z konkurenčních aplikací, protože postrádá schopnost upravovat podkladová data a nabízí méně funkcí pro zpracování dat a jejich vizualizaci, kompenzuje tyto nedostatky jasným a přímým zaměřením na jednoduché používání a přehlednou user experience (UX). [26]

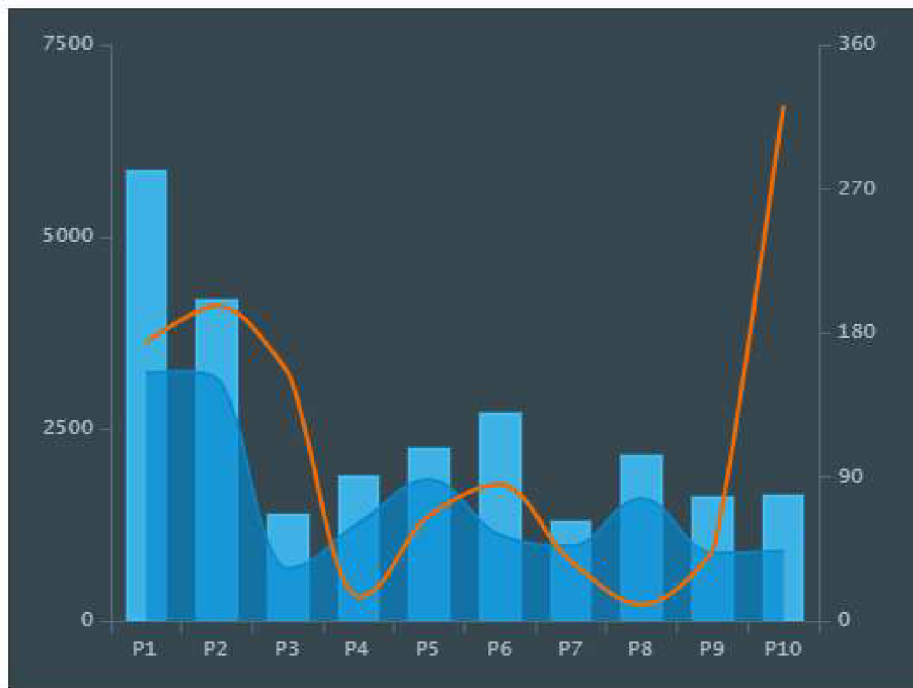
5.4 *Další nástroje pro vizualizaci dat*

AnyChart

Sami vývojáři přímo na stránkách aplikace uvádí heslo s mottem aplikace: „Vytvořeno pro vývojáře“ – to skvěle vystihuje hlavní myšlenku, která vedla ke vzniku aplikace – pro každého vývojáře by mělo být co nejsnadnější integrovat užitečné a vizuálně dobře zpracované grafy do jakéhokoli mobilního a stolního zařízení nebo webové aplikace. [27]

Mezi klienty využívající technologii AnyChart jsou přední světové firmy z nejrůznějších odvětví jako Ford, Novo Nordisk nebo Lockheed-Martin. [27]

AnyChart je vyvíjen od roku 2003. Jedná se o JavaScriptovou knihovnu, která umožňuje vizualizaci dat na více platformách současně. Nabízí interaktivní grafy, schémata a dashboardy (daly by se označit jako řídicí panel, zobrazuje data a zároveň obsahuje prvky pro ovládání). Kromě možnosti vytvořit si vlastní grafy nabízí i databázi s již vytvořenými produkty. Výhody tohoto řešení jsou pravidelné updaty a podpora ze strany vývojářů. [27]



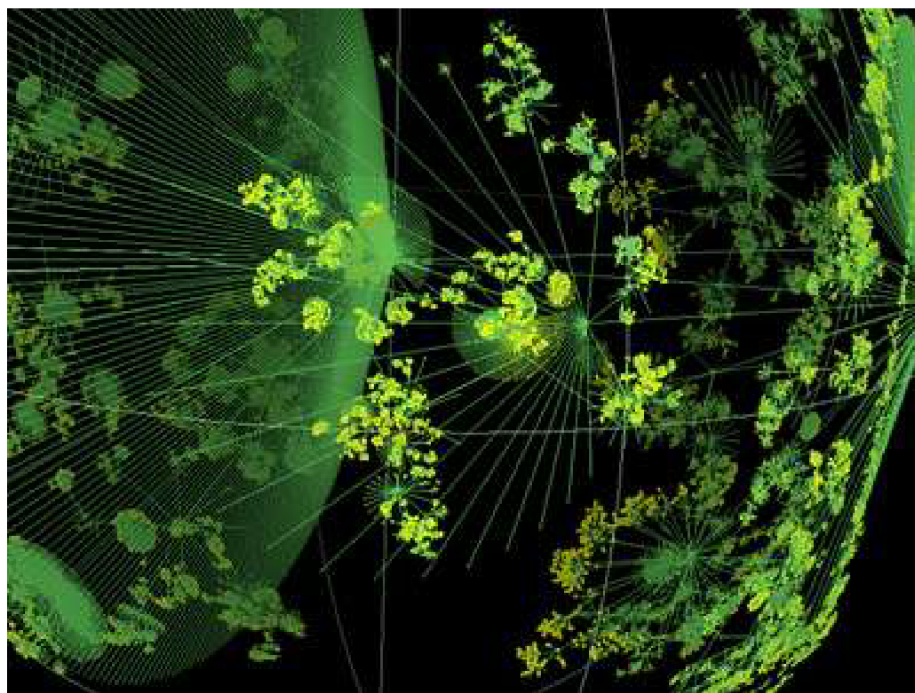
Obrázek 17 – Interaktivní graf vykreslený aplikací AnyChart [27]

Vizualizace vytvořené AnyChartem se velice snadno dají publikovat na sociálních sítích nebo exportovat do formátů .pdf, .png, .svg. Rovněž nabízí možnost exportovat data do formátu .csv nebo excelové tabulky. Výrobce uvádí, že jejich komponenta dokáže běžet na jakékoliv platformě a je schopna pracovat s jakoukoliv databází. To díky tomu, že zdrojový kód je otevřený a každý tak může modifikovat a využít aplikaci dle vlastní potřeby. [27]

Walrus

Walrus je nástroj pro vykreslování grafů ve 3D, je vyvíjen organizací CAIDA. Jeho výhoda spočívá v tom, že dokáže zpracovávat obrovská množství dat, takže umožňuje zobrazování masivních grafů (až 1 milion vrcholů). Kvůli tomu ovšem má jistá omezení, například podporuje pouze orientované grafy, nepodporuje více spojení mezi dvěma vrcholy a kvůli velikosti neumožňuje úpravy v reálném čase – musí se vykreslit znova. [28]

Tento software bohužel i přes to, že nabízí unikátní řešení a téměř neomezené možnosti vizualizace, neustál nástup novějších technologií a jeho aktivní podpora a údržba byla tak ukončena. [28]

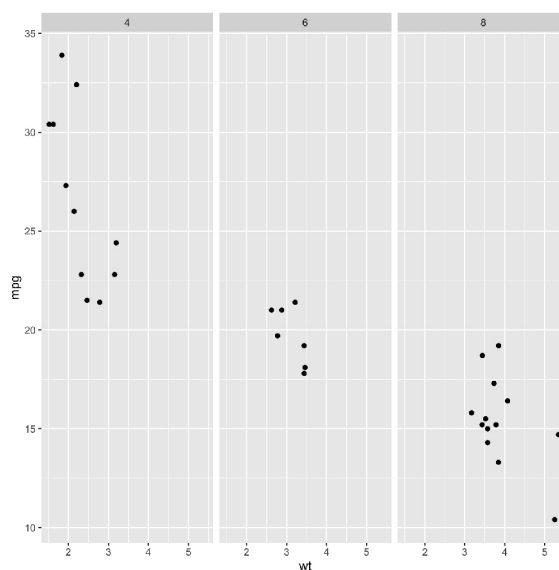


Obrázek 18 – 3D graf vykreslený aplikací Walrus [28]

5.4.1 Jazyk R

Jazyk R je programovací jazyk, který se využívá pro statistickou analýzu dat a jejich grafické zobrazení. Byl vyvinut společenstvím R CORE TEAM v roce 1993. Zajímavostí je, že jazyk R je původně implementace jazyka S, avšak pod licencí pro svobodné šíření softwaru. S jazykem R se pracuje pomocí příkazové řádky, ale existuje několik produktů, které poskytují i grafické rozhraní pro práci. Mezi nejznámější patří RStudio. Jazyk R je snadno rozšiřitelný a komunita jazyka R je velice aktivní. Většina balíčků a funkcí dostává pravidelné aktualizace. Pokročilí uživatelé mají také možnost při náročných úlohách propojit R s jinými jazyky, například C, C++, Python... [29]

ggplot2 – Jak bylo již zmíněno výše, jazyk R je díky své licenci velice snadno rozšířitelný. Mezi nejznámější balíčky, které byly pro R vytvořeny, patří knihovna ggplot2. Tato knihovna byla vytvořena přímo za účelem přinést lepší vizualizační nástroje pro R. Značně vylepšuje estetiku a celkový vizuální dojem grafických prvků. ggplot2 umožňuje vytvořit téměř všechny typy grafů, které jdou dále upravovat. Uživatel může vkládat ke grafům popisky, vkládat anotace k jednotlivým prvkům v grafu, upravovat pozadí, osy grafu a vysvětlivky. Mezi další možnosti patří rozdělení grafu na menší části, z nichž každá může být zkoumána individuálně. Na následujícím obrázku je zachycen graf vykreslený touto knihovnou. [30]



Analýza dat skladového hospodářství má význam zejména kvůli lepšímu přehledu o technickém stavu komponent a také jejich vzájemné kompatibilitě kvůli následné kompletaci počítačových sestav a jejich exportu.

Pobočka v Hradci Králové sestává z kanceláře/dílny a dvou skladů. V každém skladu jsou uložena technická zařízení, se kterými se provádí dále různé činnosti jako čištění, servis, testování, balení a expedice. Nejčastěji se jedná o počítačové komponenty: základní desky, procesory, paměti a chladiče. Nejvíce jsou ale zastoupeny ve skladu grafické karty, kterých je rozličné množství druhů. Komponenty jsou v různém technickém stavu (před/po servisu) a je proto nutné orientovat se v nich, aby bylo možné efektivně je servisovat a expedovat je koncovým zákazníkům nebo dále v rámci firmy pro další použití. Problém poslední doby je vzhledem k nepříznivému vývoji odvětví v tom, že zařízení na těžbu kryptoměn se odstavují a dávají k servisu. V malých prostorech hradecké pobočky se tak hromadí technické prostředky, které nejsou efektivně využívány.

Dalším předmětem zkoumání bude běžná kancelářská agenda jako výdaje související s provozem pobočky a analýza dat z tabulky zaměstnanců. Vzhledem k citlivé povaze některých dat nemusí data vždy přesně odrážet skutečnost, to se týká zejména jmen zaměstnanců anebo pozměněných důvěrných informací týkajících se technických prostředků (názvy, ceny).

6.1 Vizualizace dat finanční administrativy za použití Microsoft Excel

V této kapitole se práce zaměří na vizualizaci dat týkající se pohybu financí v rámci firemní pobočky. Vizualizovaná data budou s přehledem prezentovat výši výdajů, nejčastěji nakupované položky, personální agendu spojenou s financemi a podobně. Tato data se liší v některých aspektech od reálné situace.

6.1.1 Podkladová data a postup zpracování

Z tabulky jsou na první pohled zřejmá některá fakta i bez použití vizualizačních nástrojů. Vidíme kupříkladu, že se nákupy určitého typu zboží periodicky opakují každý měsíc a je také hned jasné, že naprostá většina nákupů byla nahlášena účetnímu, kromě těch z posledního měsíce. To by mohlo naznačovat, že se nákupy nahlašují vždy až na konci měsíce. Také je zřejmé, že naprostá většina nákupů byla proplacena, a to většinou Petrem. Mimo to ale z tabulky mnoho relevantních informací nevyčteme, a proto využijeme funkcí Excelu k vytvoření jednoduchých schémat a diagramů, které nám pomohou tato data lépe pochopit. Kromě zdrojových dat budou v této části uvedeny také postupy použité při zpracování. Dříve než začneme vizualizovat, položíme si několik základních otázek, na které budeme pomocí vizualizovaných dat odpovídat:

Jaká je situace? Proč chceme data vizualizovat?

Častým důvodem k vizualizaci dat je rozhodování o změnách v procesech, financování a podobně. V tomto případě je situace ohledně financí ale nastavena optimálně, a tak bude vizualizace sloužit spíše pro zpětnou kontrolu financí. Uvažujme, že měsíční rozpočet, který se v drtivé většině daří dodržet, je 2 000 Kč. Můžou ale nastat situace, že je potřeba dokoupit vybavení, jehož cena měsíční limit přesahuje (lednička, nový komponent k PC a jiné) a v takovém případě se situace řeší individuálně.

Jaká data jsou relevantní pro vizualizaci?

Vzhledem k tomu, že se zabýváme kontrolou financí, budou pro nás relevantní zejména číselná data (ceny a množství zboží), také stavy proplacení a zaúčtování nákupů, kategorie nakupovaného zboží a jména zaměstnanců, kteří nákupy a úhrady za ně prováděli. Méně relevantní už pro nás budou data uskutečnění nákupů.

Co konkrétně nám vizualizovaná data po zpracování sdělí?

Pokusíme se získat odpověď na otázky typu „Kdo nejvíce nakupuje a utrací? Kolik peněz se utratí měsíčně? Kdo proplácí platby? Které produkty jsou nejdražší

a nejčastěji nakupované?“ Hlavním předmětem zkoumání je zjistit, zda se drží zaměstnanci při nákupech v rámci limitu 2 000 Kč.

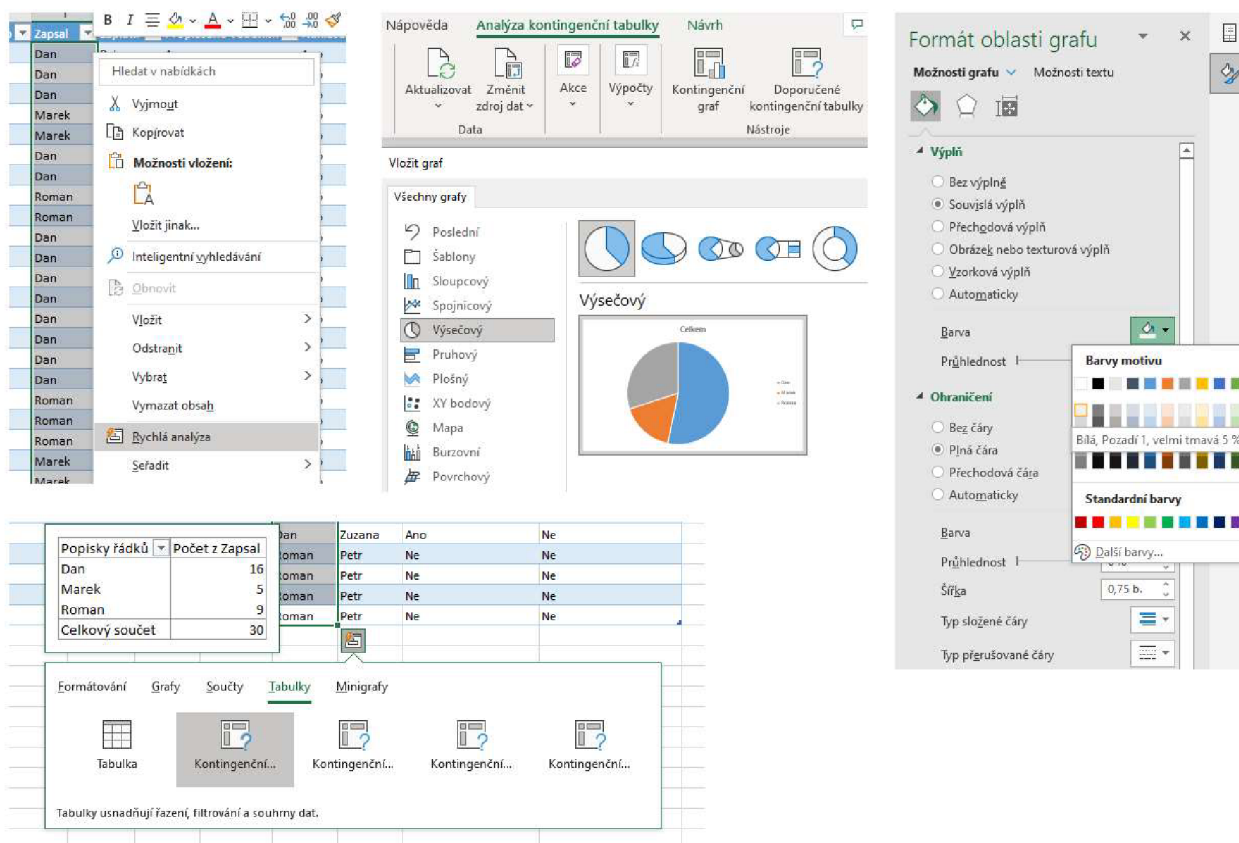
Následuje popis kroků vedoucích k vytvoření prvního diagramu. Začneme s tím, že vybereme relevantní data pro zpracování a seřadíme je v tabulce tak, aby nám Excel umožnil je dále zpracovat. Když máme data předpřipravená, začneme vizualizovat.

Vyčleníme data, která chceme vizualizovat.

Pomocí nástroje Rychlá analýza převedeme data na kontingenční tabulku

Převedeme upravená data na koláčový diagram

Upravíme diagram podle potřeby



Obrázek 20 – Použitý postup úpravy dat a tvorby diagramu

6.1.2 Výsledné vizualizace a závěry

V této části jsou prezentovány výsledky šetření. V případě dobře zpracovaných výchozích dat je vizualizace dat pomocí Excelu velice jednoduchá

a intuitivní. Proto je vhodný i pro začátečníky. V dnešní době je umění práce s Excelem součástí základní počítačové gramotnosti. Problémová data ale mohou způsobit při práci potíže. Při práci bychom se měli vyvarovat chyb, jako jsou špatně zadané hodnoty a formáty hodnot, špatně seřazená a nekompletní data a volba špatných nástrojů pro zpracování dat. Díky integraci aplikací od Microsoftu bylo možné nakopírovat schémata do práce přímo z Excelu. Tím odpadá práce s dodatečnou grafickou úpravou prezentovaných diagramů. Řešené problémy jsou uvedeny v seznamu vždy formou otázky s jasně definovatelnou odpovědí. U každé z otázek jsou prezentovány odpovědi, které vycházejí z příložených zpracovaných diagramů a schémat.

Kdo nejvíc nakupuje a kdo nejvíc utrací?

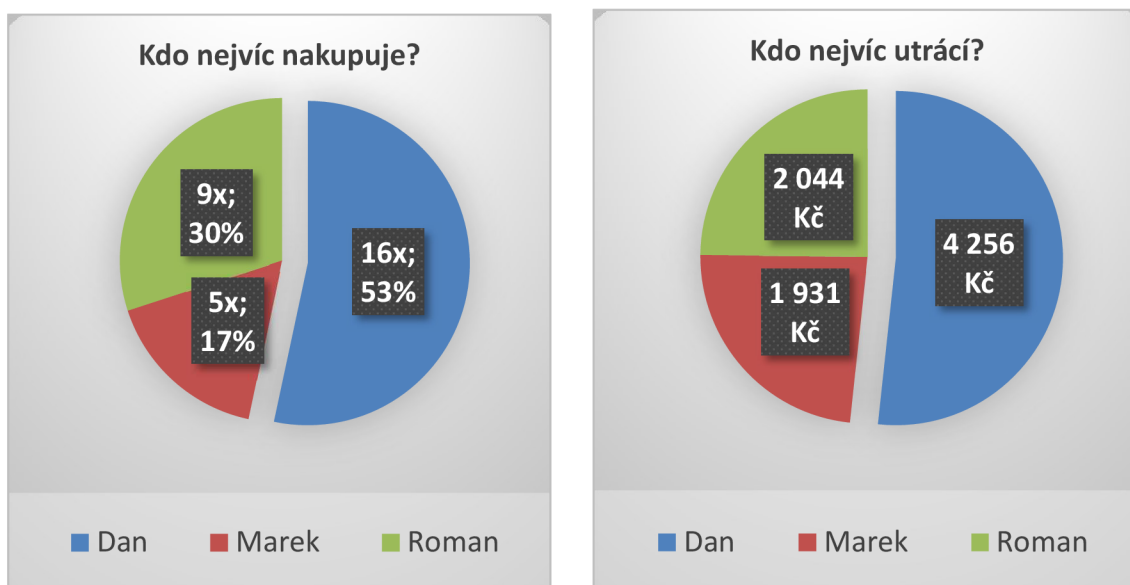


Schéma 1 – Kdo nejvíc nakupuje a utrací

Prezentované vizualizace ukazují, že nakupují pouze tři zaměstnanci pobočky. Nejvíc utrací a nakupuje Dan – zhruba polovina nákupů. Co se týče částek, tak zbylí dva zaměstnanci utrací více či méně stejně. Pro lepší přehledy byly přidány i přesně utracené částky a počty nákupů.

První diagram ukazuje modrou část o kousek větší než zbytek navzdory tomu, že má stejnou číselnou hodnotu jako zbytek diagramu. To je způsobeno tím, že je graf založený na procentech a v tomto případě pracujeme s malým celkovým počtem nákupů. V případě analýzy dat z celého roku by se graf vykreslil správně.

Za co se utrácí nejvíc a jaké je nejčastěji kupované zboží?

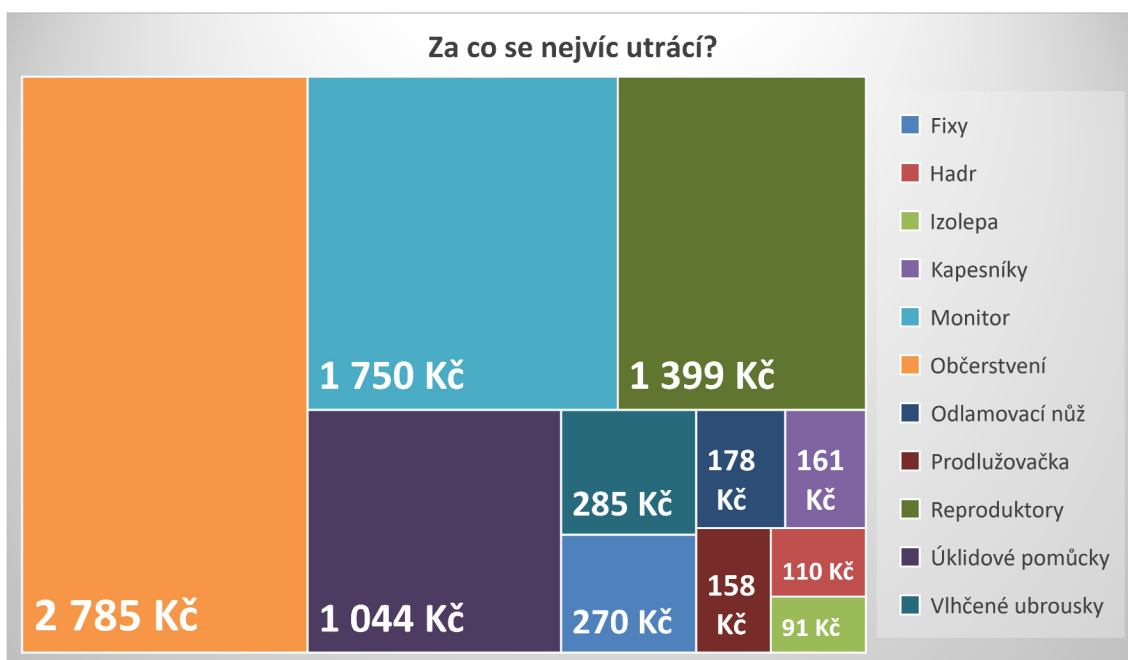


Schéma 2 – Složení celkových výdajů za období leden-červen

Pro tento přehled je možné použít různé vizualizační prostředky, ale například koláčový nebo sloupcový diagram v tomto případě nejsou vhodné, protože škála hodnot cen nakupovaného zboží má značný rozsah. Proto byla vybrána pro prezentaci dat stromová mapa – na té jsou dobře vidět barevně rozlišená pole s poměrnou velikostí z celkově utracených peněz a přesnou částkou. Na pravé straně je k dispozici legenda s popisem barevného rozlišení.

Jaké je nejčastěji nakupované zboží?

Nyní je potřeba si dodatečně určit, zda chceme zjistit celkový počet kusů zboží nakoupených za sledované období anebo zda chceme vědět, pro jaký druh zboží se nejvíce jezdí do obchodu. V prvním případě je vidět, že dost často se nakupují úklidové pomůcky, vlhčené ubrousky a kapesníky (vyšší spotřeba v zimě a na jaře). Fixy, kterých je podle grafu nejvíc, by bylo lepší uvádět s počtem kusů 1 a celkovou částkou spíše než s počtem kusů. V druhém případě je hned z grafu jasné, že nejčastěji se nakupuje občerstvení pro zaměstnance a úklidové pomůcky. Ostatní věci související s činností kanceláře jsou v nákupech zastoupeny rovnoměrně.

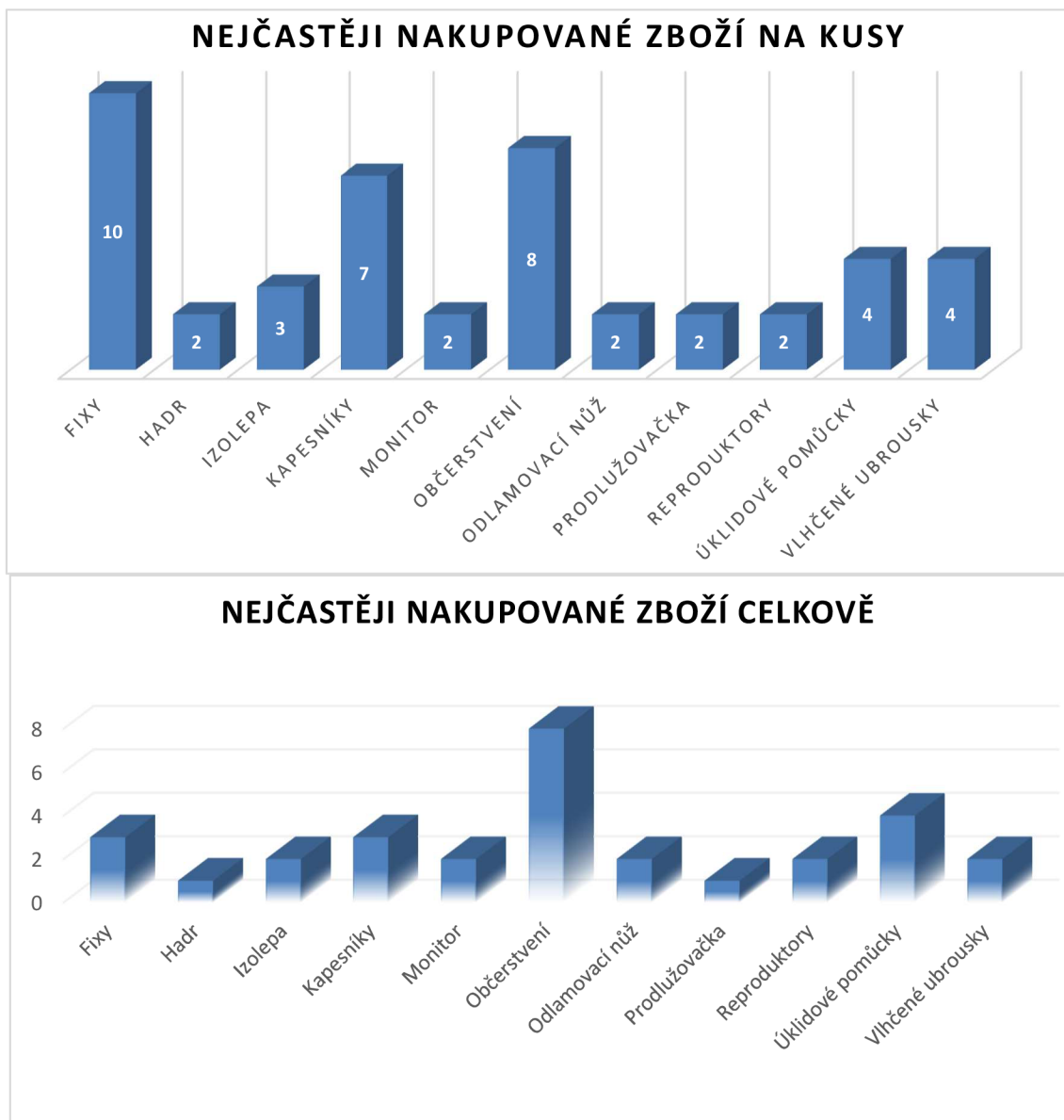


Schéma 3 – Nejčastěji nakupované zboží v kanceláři

Kdo nejvíc proplácí nákupy?

K zodpovězení této otázky byl použitý horizontální skupinový sloupcový graf. Pro porovnání jsou modře opět uvedené i počty nákupů podle zaměstnanců. Červeně jsou vidět samotné proplacené nákupy lidmi z vedení. Z diagramu je patrné, že financování provozu kanceláře zajišťuje převážně Petr. Také je patrné, že některé nákupy ještě nebyly zaměstnancům proplaceny.

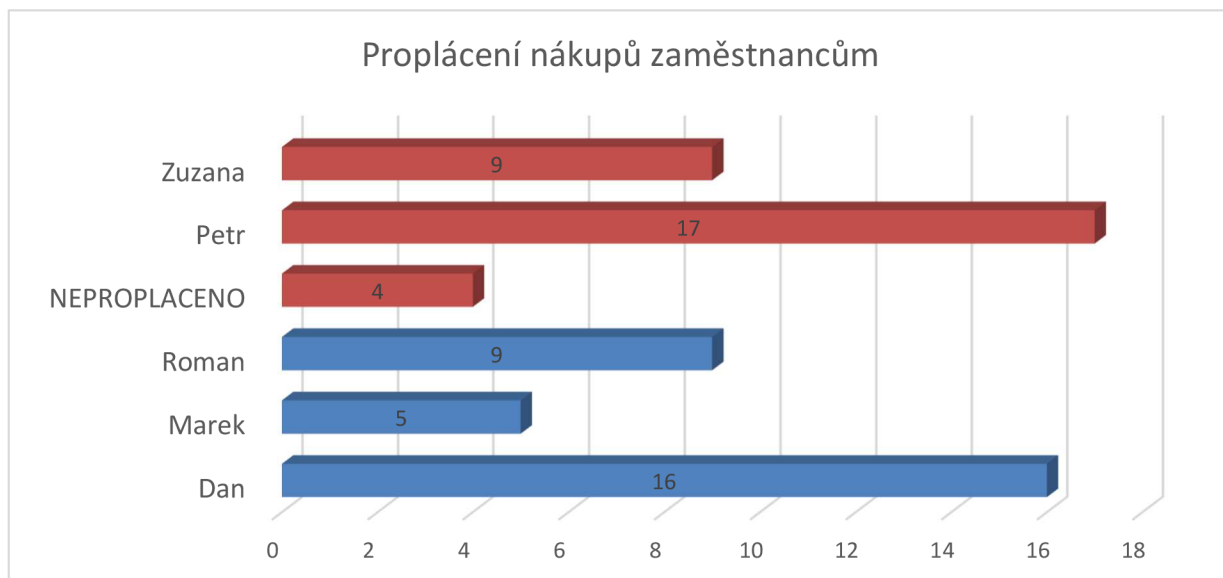


Schéma 4 – Proplacené a uskutečněné nákupy

Který měsíc se utrácelo nejvíce a byl dodržen měsíční finanční limit?

Toto jsou ty nejdůležitější otázky z pohledu zpětné kontroly financí v kanceláři. Z grafu zjistíme, že nejvíce se utrácelo v březnu, zejména na to měl vliv nákup reproduktorů k PC. Částka byla vysoká a vedení rozhodlo, že její proplacení se rozdělí i do následujícího měsíce. Jinak jsou výdaje spojené s provozem kanceláře většinou v rozmezí 1 000 – 1 500 Kč. Rovněž je jasně vidět, že výdaje nepřesáhly v žádný měsíc limit 2 000 Kč.

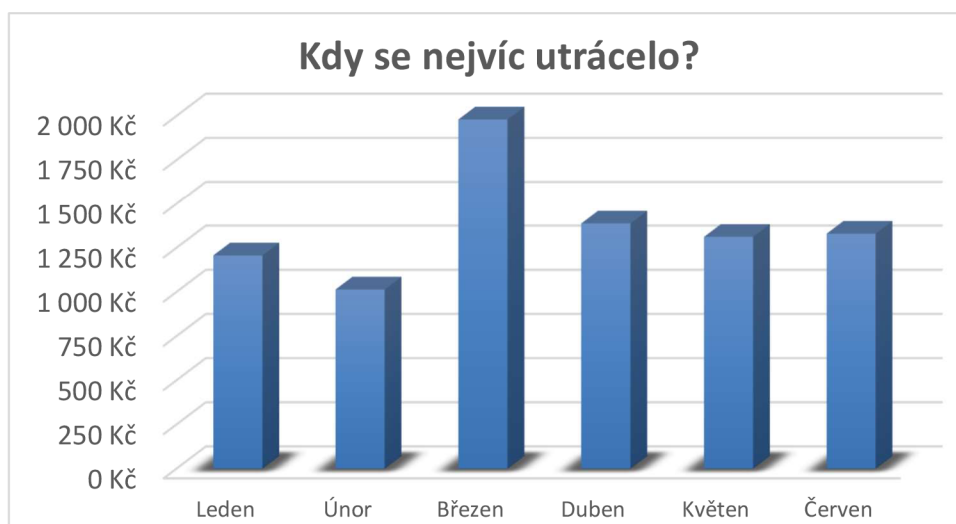


Schéma 5 – Kontrola financí v čase a dodržování limitu

Následuje tabulka s výchozími daty:

Měsíc	Datum zadání	Datum nákupu	Položka	Počet ks	Cena/ks	Celkem	Koupeno	Zapsal	Zaplatil	Proplaceno vedením	Nahlášeno na účtárnu
Leden	06.01.2022	06.01.2022	Občerstvení	1	228	228	Ano	Dan	Petr	Ano	Ano
Leden	06.01.2022	06.01.2022	Úklidové pomůcky	1	350	350	Ano	Dan	Petr	Ano	Ano
Leden	06.01.2022	06.01.2022	Fixy	4	30	120	Ano	Dan	Petr	Ano	Ano
Leden	26.01.2022	26.01.2022	Občerstvení	1	452	452	Ano	Marek	Petr	Ano	Ano
Leden	26.01.2022	26.01.2022	Kapesníky	3	19	57	Ano	Marek	Petr	Ano	Ano
Únor	16.02.2022	16.02.2022	Vlhčené ubrousky	3	75	225	Ano	Dan	Zuzana	Ano	Ano
Únor	16.02.2022	16.02.2022	Izolepa	2	26	52	Ano	Dan	Zuzana	Ano	Ano
Únor	21.02.2022	21.02.2022	Občerstvení	1	693	693	Ano	Roman	Petr	Ano	Ano
Únor	21.02.2022	21.02.2022	Kapesníky	2	22	44	Ano	Roman	Petr	Ano	Ano
Březen	03.03.2022	03.03.2022	Reproduktory	1	1000	1000	Ano	Dan	Petr	Ano	Ano
Březen	03.03.2022	03.03.2022	Prodlužovačka	2	79	158	Ano	Dan	Petr	Ano	Ano
Březen	10.03.2022	10.03.2022	Úklidové pomůcky	1	156	156	Ano	Dan	Petr	Ano	Ano
Březen	10.03.2022	10.03.2022	Občerstvení	1	311	311	Ano	Dan	Petr	Ano	Ano
Březen	11.03.2022	11.03.2022	Fixy	4	29	116	Ano	Dan	Petr	Ano	Ano
Březen	11.03.2022	11.03.2022	Hadr	2	55	110	Ano	Dan	Petr	Ano	Ano
Březen	19.03.2022	19.03.2022	Odlamovací nůž	1	129	129	Ano	Dan	Petr	Ano	Ano
Duben	01.04.2022	16.04.2022	Reproduktory	1	399	399	Ano	Dan	Zuzana	Ano	Ano
Duben	20.04.2022	20.04.2022	Občerstvení	1	430	430	Ano	Roman	Zuzana	Ano	Ano
Duben	20.04.2022	20.04.2022	Úklidové pomůcky	1	298	298	Ano	Roman	Zuzana	Ano	Ano
Duben	20.04.2022	20.04.2022	Izolepa	1	39	39	Ano	Roman	Zuzana	Ano	Ano
Duben	30.04.2022	30.04.2022	Fixy	2	17	34	Ano	Marek	Petr	Ano	Ano
Duben	30.04.2022	30.04.2022	Občerstvení	1	188	188	Ano	Marek	Petr	Ano	Ano
Květen	23.05.2022	23.05.2022	Občerstvení	1	112	112	Ano	Dan	Petr	Ano	Ano
Květen	31.05.2022	31.05.2022	Monitor	1	1200	1200	Ano	Marek	Zuzana	Ano	Ano
Červen	03.06.2022	03.06.2022	Monitor	1	550	550	Ano	Dan	Zuzana	Ano	Ne
Červen	03.06.2022	03.06.2022	Úklidové pomůcky	1	240	240	Ano	Dan	Zuzana	Ano	Ne
Červen	14.06.2022	14.06.2022	Občerstvení	1	371	371	Ano	Roman	NEPROPLACENO	Ne	Ne
Červen	15.06.2022	15.06.2022	Kapesníky	2	30	60	Ano	Roman	NEPROPLACENO	Ne	Ne
Červen	15.06.2022	15.06.2022	Odlamovací nůž	1	49	49	Ano	Roman	NEPROPLACENO	Ne	Ne
Červen	15.06.2022	15.06.2022	Vlhčené ubrousky	1	60	60	Ano	Roman	NEPROPLACENO	Ne	Ne

Tabulka 1 – Přehled financí v kanceláři za období leden-červen

6.2 Vizualizace dat personální agendy za použití Google Data studio

V této kapitole budou za pomoci vizuálně zpracovaných dat zodpovězeny otázky týkající se personální agendy. Jako v předchozí části, i v této budou nejprve uvedeny důvody, proč je potřeba data vizualizovat. V další části budou opět konkrétní otázky, které si je potřeba zodpovědět. Na konci kapitoly jsou uvedena výchozí data.

Práce s aplikací Google Data Studio je oproti Excelu náročnější, uživatelé by měli mít již nějakou zkušenost s vizualizačními nástroji. I přesto je ale aplikace vcelku intuitivní a uživatel, který již pracoval například s Excelem, snadno pochopí základní principy a funkcionalitu a aplikace. Při vybrání vhodných dat umožňuje provádět vizualizace velice snadno. Aplikace obsahuje také kurz pro nové uživatele, který vysvětlí základy. Nesporná výhoda aplikace Google Data Studio je ta, že je dostupný zdarma každému s účtem Google a také to, že veškeré renderování vizualizací a zpracovávání dat využívá výkonu serverů, na kterých aplikace běží, proto je práce s ní plynulá. Naopak velkou nevýhodou je nutnost neustálého připojení k internetu. Na druhou stranu v případě, že připojení vypadne, je vše uloženo a při opětovném připojení lze pokračovat v práci.

Google Data Studio je užitečným nástrojem pro vizualizaci dat z tabulek. Využívá se ale hlavně pro analýzu velkých dat, jejichž poskytovatelem je Google (Google Ads, Youtube Analytics) a proto je zaměřena spíše na tuto funkcionalitu.

6.2.1 Podkladová data a postup zpracování

Nejprve musíme načíst zdrojová data do aplikace. Budeme vycházet ze stejného .xlsx souboru s daty, jako ve zbytku práce, ale v případě této aplikace musíme nejprve nahrát excelový soubor na Disk Google a využít možnost převedení na formát tabulky podporovaný aplikací Google Data Studio. Jak již bylo uvedeno v teoretické části, je tato aplikace navržena především jako nástroj, který má za úkol zpracovávat především data shromážděná Googlem za využití jeho ostatních platforem a aplikací. Import dat ze služeb Google Analytics a Google Ads je tak velice

snadný, také je hned mezi prvními podporovanými zdroji dat možnost importu z tabulek Google (pro Excel nutná zmíněná úprava) anebo přímo z .csv souboru.

Poté, co zvolíme zdroj dat, která chceme vizualizovat, nám aplikace otevře hlavní pracovní plochu, kde na pravé straně okna máme zobrazené zdrojové tabulky a v horní části okna jsou dostupné nástroje pro dodatečnou úpravu dat a dostupné vizualizační nástroje. Ještě, než začneme s vizualizací, zodpovíme si základní otázky:

Jaká je situace? Proč chceme data vizualizovat?

V tomto případě by se nám hodilo rozdělit zaměstnance do různých skupin založených na typu smlouvy, poloze bydliště, dojezdové vzdálenosti do práce a dalších faktorech jako jsou přístup k financím nebo vlastnictví řidičského oprávnění.

Jaká data jsou relevantní pro vizualizaci?

V rámci zodpídaných otázek budou pro nás z tabulky zaměstnanců relevantní zejména data typu druh úvazku, status studenta, majitel ŘP, přístup k financím, bydliště v HK a vzdálenost do práce. Naopak nerelevantní budou jména zaměstnanců a pozice, na kterých pracují. Již na první pohled je zřejmé, že zaměstnanci pobočky v HK jsou především technici – brigádníci zabývající se čištěním a tříděním komponent.

Co konkrétně nám vizualizovaná data po zpracování sdělí?

Vzhledem k přeplněnosti skladu v HK je nutné odvážet zpracované komponenty jinam, chceme tedy znát odpověď na otázku „kolik zaměstnanců je majitelem ŘP sk. B?“ Získáme také informace o tom, kolik zaměstnanců v současné době studuje, kolik jich bydlí v místě zaměstnání a kolik jich musí dojíždět.

6.2.2 Výsledné vizualizace a závěry

V tomto případě budou schémata a vizualizace z aplikace vkládána do práce formou obrázků. Vždy bude uvedena konkrétní otázka a diagram, ze kterého lze vyčíst odpověď. Také budou uvedeny komentáře s vyvozenými závěry a doporučeními vycházející z vizualizovaných dat.

Jaký je podíl zaměstnanců schopných převážet komponenty?

Jak již bylo zmíněno, v současnosti je největším problémem hromadění komponent ve skladu. Firma má sice uzavřenou smlouvu s profesionálním řidičem, ale může se stát, že jeho služby z nějakého důvodu nebudou dostupné (nemoc, cesta do zahraničí, porucha vozu) a proto by bylo dobré vědět, kolik zaměstnanců by mohlo v případě potřeby komponenty odvézt. V rámci bezpečnosti práce by pro tuto činnost byli vybíráni zejména zaměstnanci pracující na HPP. Google Data Studio nám umožní navolit do kontingenční tabulky vybrané sloupce a v následně vygenerovaném grafu podle podmínek zobrazit počet zaměstnanců s HPP, kteří jsou současně majiteli řidičského oprávnění.

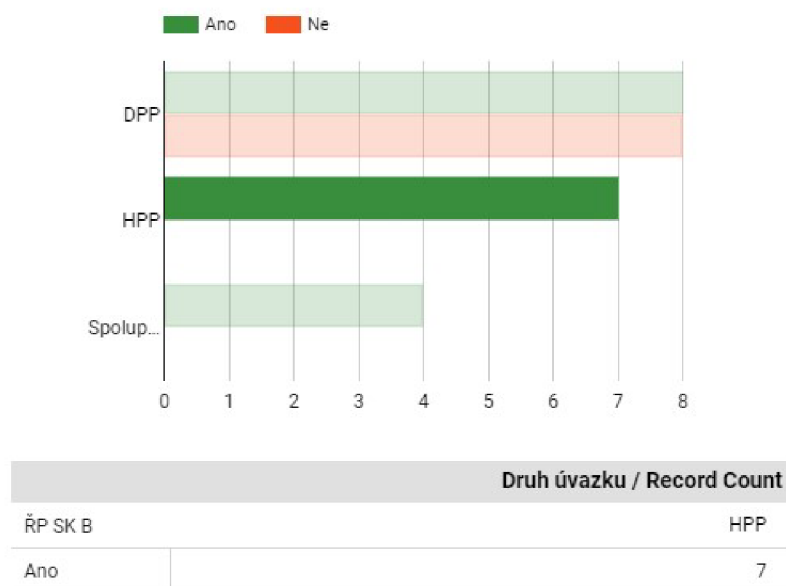
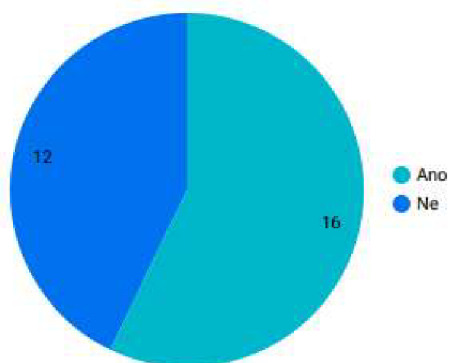
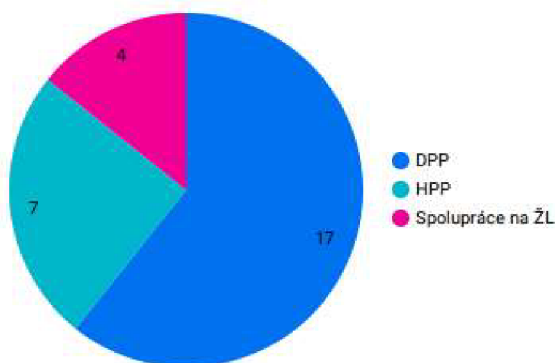


Schéma 6 – Počet vyhovujících řidičů

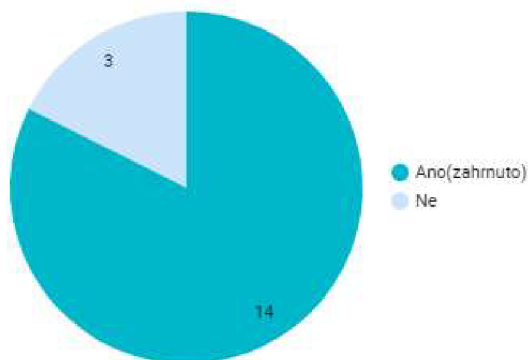
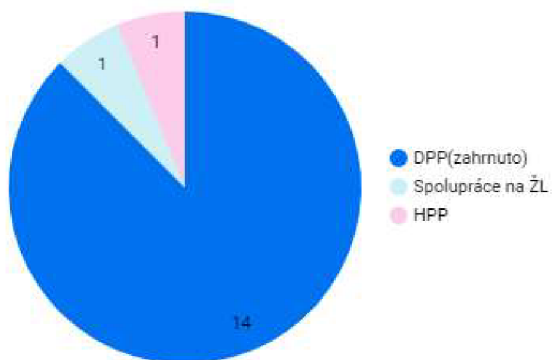
Jaké je rozdělení typu pracovního úvazku mezi zaměstnanci a kolik jich pracuje při studiu?

V případě pobočky v HK má většina zaměstnanců s firmou uzavřenou dohodu o provedení práce (DPP), to je dané tím, že převážná většina jsou studenti SŠ a VŠ, kteří pracují jako brigádníci a provádí většinou jen jednoduché činnosti nevyžadující značné technické znalosti a dovednosti. Informace o tom, zda je zaměstnanec studentem, jsou důležité pro účetní oddělení kvůli daním a výplatám, v případě plánování pracovní doby bude počítáno s těmito zaměstnanci spíše na odpolední nebo večerní práci. Zbytek zaměstnanců je veden buď na hlavní pracovní

poměr (HPP), zejména management a vedoucí technický personál, nebo spolupracují externě na živnostenský list. V tomto případě Google data studio dovoluje vytvořit propojenou tabulku, pomocí které můžeme zjistit přesný počet daných úvazků a následně vybrat pouze studenty pracující na DPP.



Druh úvazku / Record Count			
Student	DPP	HPP	Spolupráce na ŽL
Ano	14	1	1
Ne	3	6	3



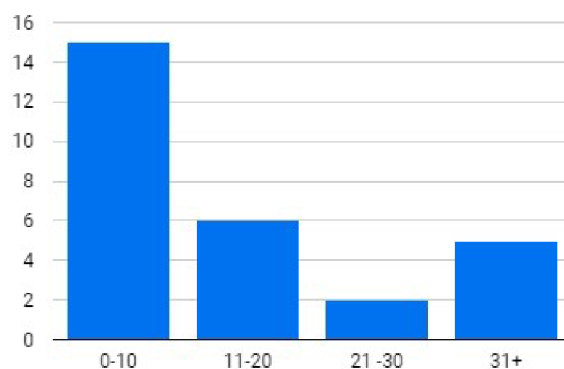
Druh úvazku / Record Count	
Student	DPP
Ano	14

Schéma 7 – Druhy úvazků zaměstnanců

Bydlí zaměstnanec v HK, dojíždí, případně z jaké vzdálenosti?

Tyto informace jsou opět klíčové v případě, že by bylo nutno plánovat práci dopředu. Pracovníkům z větší vzdálenosti (21-30 km, 31+ km) by bylo vhodné plánovat delší směny (8 hodin / 12 hodin) v rámci jednoho dne, nebo několika dní po sobě, aby se jim cesta do HK vyplatila. Naopak zaměstnanci bydlící v HK by mohli

chodit na kratší směny (okolo 4 hodin denně) rozvrhnutých do více dnů. Je vidět, že naprostá většina zaměstnanců je z HK (0-10), nebo blízkého okolí (do 20 km), zejména se jedná o techniky, kteří jsou klíčoví pro zpracování a třídění komponent. Celkem se jedná o 21 z celkových 28 zaměstnanců. Vedení, účetní, personální a podobný personál pracuje v Praze na firemní centrále, proto jsou uvedeni v kategorii 31+ km, těch by se ale případné plánování směn netýkalo.

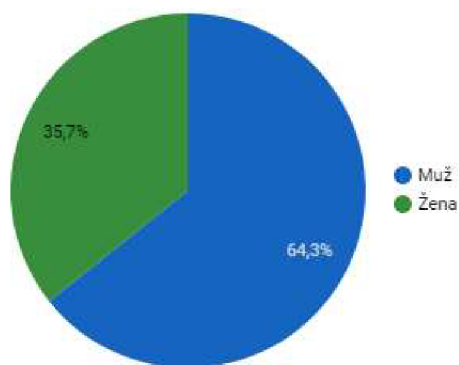


Vzdálenost v km	Record Count
0-10	15
11-20	6
21-30	2
31+	5

Schéma 8 – Dojezdová vzdálenost zaměstnanců do práce

Jaký je poměr mužů a žen mezi zaměstnanci?

Tato otázka byla položena především proto, že odpověď na ní může být překvapivá. Zejména proto, že v technických oborech souvisejících s IT nebývají ženy tak často zastoupeny. Tím spíše je překvapivé, že ženy mezi zaměstnanci mají podíl více než z 1/3. A to navzdory tomu, že součástí práce je často i manipulace s těžkými břemeny ve skladu. Dalo by se tak říci, že situace v této firmě odráží trend poslední doby, kdy stále více žen volí práci v technických oborech, zejména IT.



	Pohlaví	Record Count
1.	Muž	18
2.	Žena	10

Schéma 9 – Zastoupení žen a mužů mezi zaměstnanci

Následuje výchozí tabulka s daty zaměstnanců:

Jméno	Pohlaví	Druh úvazku	Student	ŘP SK B	Pozice	Správa financí	Bydliště v HK	Vzálenost v km	Hodiny
Andrea	Žena	DPP	Ano	Ano	Technik	Ne	Ano	0-10	20
Barbora H.	Žena	DPP	Ano	Ano	Technik	Ne	Ano	0-10	20
Barbora O.	Žena	DPP	Ano	Ano	Technik	Ne	Ne	11-20	16
Dan K.	Muž	HPP	Ne	Ano	Vedoucí technik	Ano	Ano	0-10	135
Dan O.	Muž	DPP	Ano	Ne	Vedoucí technik	Ano	Ano	0-10	81
Jakub	Muž	DPP	Ano	Ne	Technik	Ne	Ano	0-10	22
Jan M.	Muž	Spolupráce na ŽL	Ne	Ano	Technik	Ne	Ne	11-20	55
Jan S.	Muž	HPP	Ne	Ano	Účetní	Ne	Ne	31+	152
Jiří	Muž	DPP	Ano	Ano	Technik	Ne	Ano	0-10	0
Kateřina K.	Žena	DPP	Ano	Ano	Technik	Ne	Ne	11-20	0
Kateřina T.	Žena	DPP	Ano	Ne	Technik	Ne	Ano	0-10	16
Lenka	Žena	HPP	Ne	Ano	Manažer	Ano	Ano	0-10	130
Luboš	Muž	DPP	Ne	Ano	Technik	Ne	Ne	11-20	96
Lukáš	Muž	DPP	Ano	Ne	Technik	Ne	Ano	0-10	0
Marek	Muž	HPP	Ne	Ano	Software adminis	Ano	Ano	0-10	120
Marie	Žena	HPP	Ano	Ano	Personální	Ano	Ne	31+	148
Matěj J.	Muž	DPP	Ano	Ne	Technik	Ne	Ne	11-20	0
Matěj R.	Muž	DPP	Ano	Ano	Technik	Ne	Ano	0-10	38
Matyáš	Muž	Spolupráce na ŽL	Ne	Ano	Řidič	Ano	Ne	31+	60
Naomi	Žena	DPP	Ne	Ne	Technik	Ne	Ne	21-30	63
Patrik	Muž	DPP	Ano	Ano	Technik	Ne	Ne	21-30	63
Petr B.	Muž	HPP	Ne	Ano	Manažer	Ano	Ne	31+	158
Petr J.	Muž	Spolupráce na ŽL	Ne	Ano	Vedoucí technik	Ano	Ano	0-10	112
Roman	Muž	DPP	Ne	Ano	Vedoucí technik	Ano	Ano	0-10	84
Sebastian	Muž	DPP	Ano	Ne	Technik	Ne	Ano	0-10	12
Standa	Muž	Spolupráce na ŽL	Ano	Ano	Technik	Ne	Ne	31+	36
Tereza	Žena	HPP	Ne	Ano	Technik	Ne	Ne	11-20	151
Zuzana	Žena	DPP	Ano	Ne	Technik	Ne	Ano	0-10	0

Tabulka 2 – Data zaměstnanců pobočky

6.3 Vizualizace dat skladového hospodářství za použití Microsoft Power BI Desktop

Nyní se práce zaměří na vyřešení situace skladového hospodářství. Bude využita aktuální verze aplikace MS Power BI Desktop. Tato kapitola je nejobsáhlejší a řeší problémy s nedostatkem místa a nepřehledností v komponentech. Primární cíle jsou v tuto chvíli dva:

Zprv je potřeba přebrat a otestovat základní desky. Desek je v současnosti velké množství hlavně v prostorech dílny a kanceláře a kvůli tomu je stísněný prostor na práci. Funkční desky budou následně osazeny procesory a RAM. Některé takovéto sestavy budou odeslány zpět do procesu a jiné budou určeny k prodeji zákazníkům.

Dalším, dlouhodobějším cílem je testování GPU, které jsou buď před nebo po servisu. Podle dále uvedených kritérií bude rozhodnuto o servisu nebo likvidaci/prodeji nefunkčních kusů. Funkční budou následně odesílány zpět do procesu nebo rozprodány.

6.3.1 Podkladová data a postup zpracování

Aplikace Power BI nabízí import dat z .xlsx a .csv souborů hned jako jednu z prvních možností. Mimo to je také možnost propojení nejrůznějších SQL databází a také analytických online služeb (Power BI Service, Google Analytics, Microsoft Azure, Dynamics 365).

V tomto případě vycházíme opět vycházíme z Excelových tabulek, ale data jsou obsáhlejší. Je to zapříčiněno tím, že vycházíme z dat o různých komponentech. Konkrétně budeme analyzovat grafické karty (GPU), procesory (CPU), paměti RAM a základní desky (Motherboard). V tomto případě máme také tabulku s nefunkčními a poškozenými komponenty. Jako v předchozích kapitolách budou zodpovězeny otázky, které definují základní důvody a cíle vizualizací.

Jaká je situace? Proč chceme data vizualizovat?

Akutní situace související s nedostatkem místa ke skladování komponent zapříčinila, že v současnosti jsou komponenty skladovány v neideálních

podmínkách (krabice, regály) a to může mít za důsledek poškození komponent. Také v některých případech dokonce nemůže probíhat více pracovních činností najednou (osazování sestav x testování GPU). Je potřeba jasně vyobrazit celkovou situaci ohledně komponent a na základě zjištěných dat rozhodnout o dalších pracovních postupech.

Jaká data jsou relevantní pro vizualizaci?

V tomto případě budou hlavní počty kusů jednotlivých komponent. U GPU budeme zjišťovat hlavně typ a výrobce grafického čipu a stav údržby. U základních desek jsou klíčové údaje o patici procesoru a typu paměti RAM. U nefunkčních komponent se budeme zabývat druhem poškození. Méně relevantní budou v některých případech konkrétní typy a názvy komponent. Kvůli důvěrnosti informací nejsou uváděna data jako cena komponent a jména pracovníků, kteří s nimi manipulovali.

Co konkrétně nám vizualizovaná data po zpracování sdělí?

Je potřeba zjistit stav komponent v rámci údržby, celkové počty a vzájemnou kompatibilitu komponent. Kompatibilita je důležitá pro osazování sestav určených k prodeji nebo opětovnému nasazení do procesu. U grafických karet chceme zjistit především počty a poměry množství karet s různými čipy. Také chceme zjistit nejčastější druhy poškození komponent.

6.3.2 Výsledné vizualizace a závěry

V tomto případě jsou opět vloženy grafy a diagramy jako obrázky. Power BI mimo jiné umožňuje také filtrovat data přímo v grafech, upravovat vizuální stránku a průběžně měnit zdrojová data bez nutnosti vytvářet další grafy. V tomto případě nejprve zjistíme typy komponent a poté je rozdělíme do skupin podle využitelnosti.

Jaké typy GPU a typy GPU čipů jsou na skladě?

Z analýzy vyplývá, že většina karet je s čipem od výrobce AMD, asi 2/3. AMD GPU byly v době pořizování cenově dostupnější. Také je vidět, že největší podíl mezi grafickými čipy mají RX570, RX580 od AMD a nVidia GTX1060. také jsou výrazně zastoupeny karty s čipy RX470, RX480 a novější GTX1070 a GTX1080.

Tyto typy karet byly využívány pro těžbu v době „bitcoin boomu“ po celém světě. Z výrobců je nejvíce zastoupena firma Sapphire Technology, která již od vzniku vyrábí výhradně karty s čipy AMD, po ní MSI a Asus (počtem GPU je ale nejvíce zastoupena firma XFX, viz další vizualizace) Vzhledem k momentálnímu požadavku, aby se navrátily karty zpět do procesu, budou otestovány primárně právě karty s čipy RX570, RX580 a také RX470.

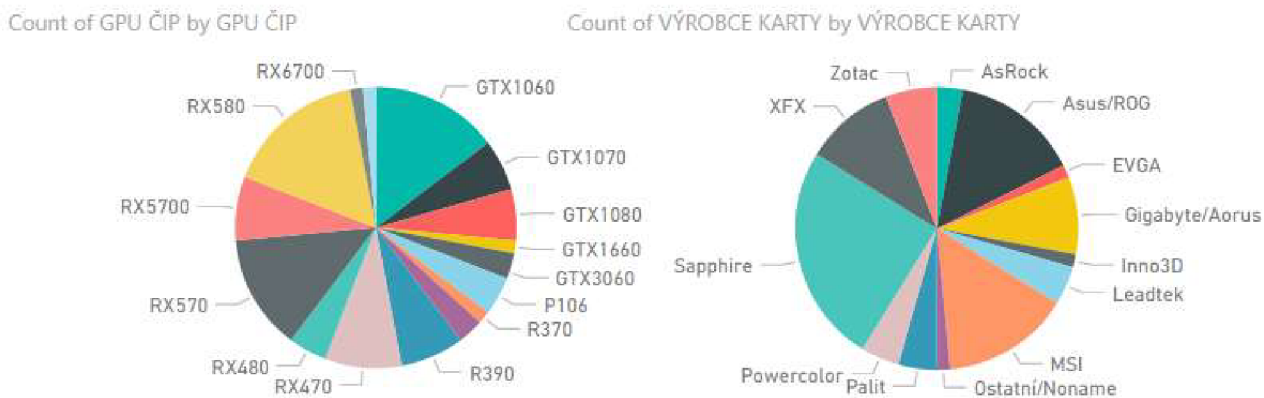


Schéma 10 – Typy a výrobci GPU na skladě

Jaké jsou konkrétní počty karet v servisu?

Jak již bylo zmíněno, pobočka firmy v HK se zabývá servisem PC komponent. Komponenty jsou dodávány ze tří hal. Vzhledem k současné situaci se budeme zabývat daty z haly 1 a 2. V případě karet z obou hal budou testovány nejprve karty s čipy od AMD, kvůli současné konfiguraci testovacích sestav. Poté budou testovací sestavy přeinstalovány na testování karet nVidia.

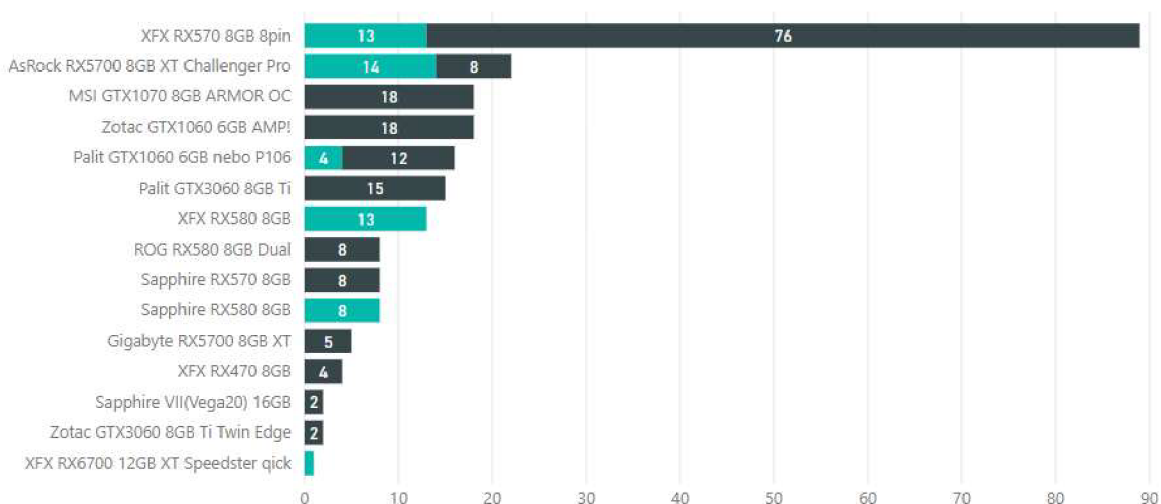
Momentální úkol je otestování a případné opravy karet z haly 1, které mají soukromého majitele, firma je má pouze „zapůjčené“ a majitel požaduje jejich opětovné zapojení do procesu. Z následujících schémat je patrné, že z haly 1 je nejvíce karet s pamětí 8 GB (celkem 192 ks), ty mají být testovány primárně. Horní graf vyobrazuje typy karet, nejvíce je karet XFX RX570 8GB 8pin, které se budou testovat jako první. Poté se otestují ostatní RX570 a RX580 karty. Hodnoty v grafu jsou děleny na dvě skupiny REFLOW ANO X NE, reflow je specifický druh servisu, kdy se za působení tepla opětovně usazují tranzistory a čipy zpět na desku s tištěnými spoji. Nejprve se testují karty bez reflow a pokud jsou nefunkční proběhne servis. Je ověřeno, že se po reflow v nezanedbatelném počtu případů

problém s kartou vyřeší. Karty po reflow jsou následně opět otestovány a spolu s funkčními kusy odesílány zpět do haly 1. Nefunkční se balí buď k prodeji nebo se likvidují (mechanické poškození).

Po dokončení karet z haly 1 se budou testovat karty z haly 2, primárně ty s čipem RX470, které budou následně také zapojeny do procesu. Z grafu je vidět, že i v jako minulém případě je nejvíce karet s čipem RX470 především od výrobce XFX, konkrétně 13 ks, dalších 13 ks je od jiných výrobců. Jako další budou testovány karty s čipy RX480, RX570 a RX580 s kapacitou paměti 8 GB. 4 GB verze těchto karet budou zřejmě určeny k prodeji. Pro potřebu tohoto grafu byly vynechány údaje o kartách s čipy nVidia GTX. Příklad filtru je vidět vedle grafu.

POČET KS by TYP GPU and REFLOW

REFLOW ● ANO ● NE



POČET KS by PAMĚŤ GB and PAMĚŤ GB

PAMĚŤ GB ● 6 ● 8 ● 12 ● 16

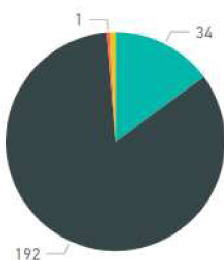


Schéma 11 – GPU z haly 1 v servisu

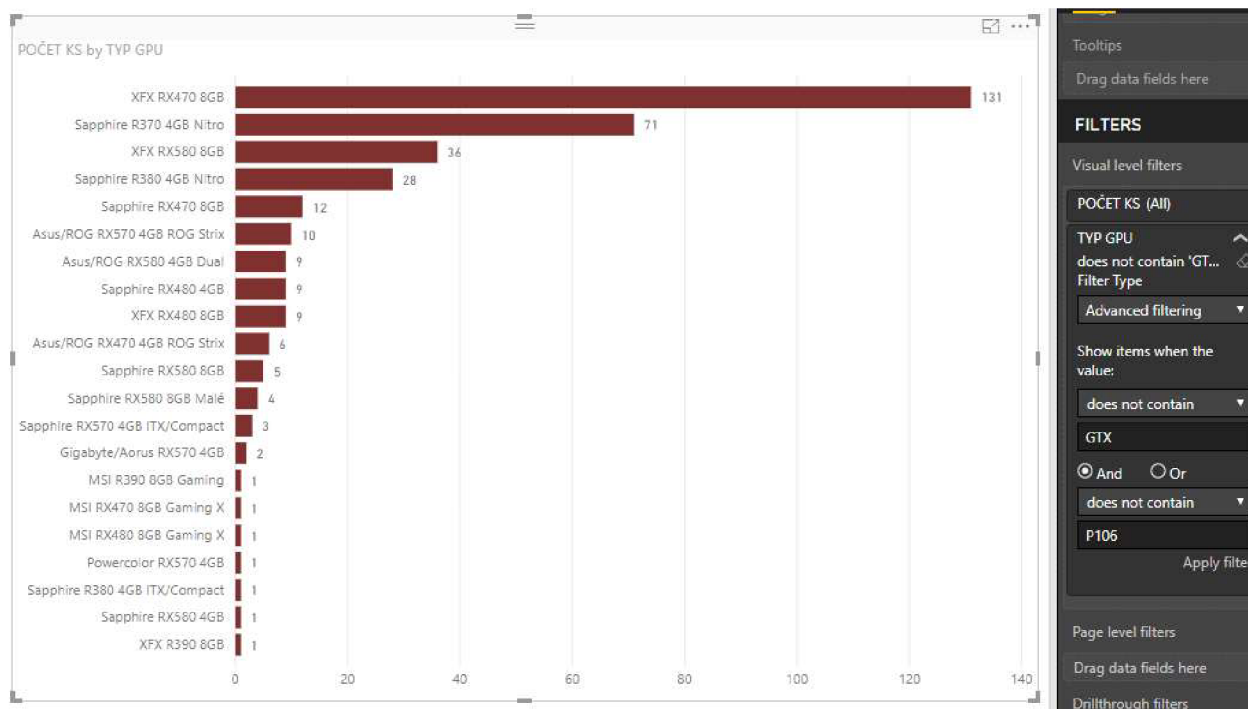
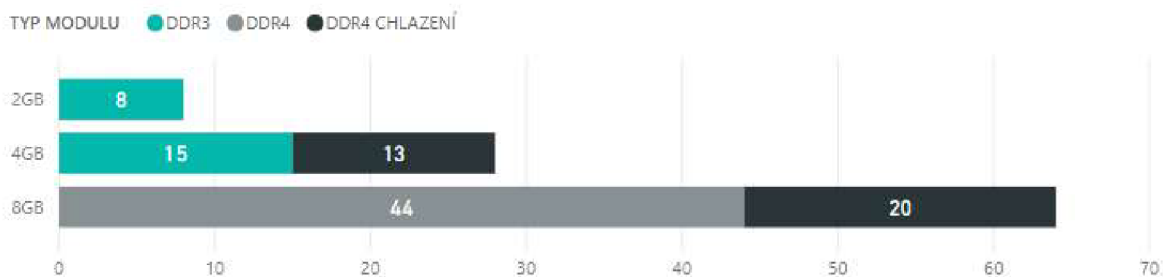


Schéma 12 – GPU z haly 2 v servisu a podmínka filtru v Power BI

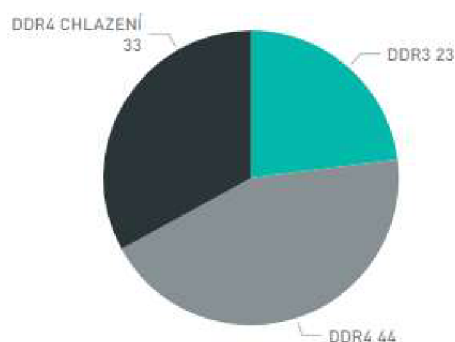
Jaký je stav komponent potřebných k tvorbě PC sestav?

V tomto případě se jedná o CPU, RAM a samotné základní desky. Ne všechny komponenty jsou kompatibilní (jiné generace pamětí, patice CPU) a proto je potřeba zjistit kolik je možné osadit desek a odeslat je dále do procesu či k prodeji. Sestavy s procesory AMD R5 a vybrané sestavy se staršími procesory Intel Core i3 jsou určeny k nasazení zpět do procesu, ostatní desky s procesory Intel Celeron/Pentium jsou určeny k prodeji jako low-end PC. V následujících grafech je vidět počet kusů jednotlivých CPU, pamětí a desek. CPU a RAM byly otestovány na jiných sestavách a položky v následujícím grafu jsou tedy všechny funkční.

POČET KS by KAPACITA and TYP MODULU



POČET KS by TYP MODULU



POČET KS by TYP CPU and PATICE

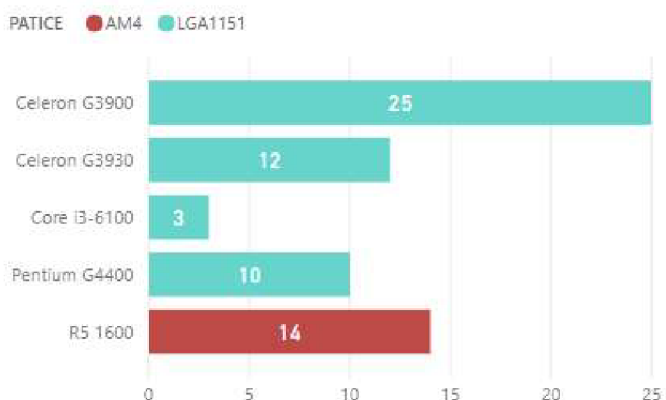


Schéma 13 – RAM a CPU určené k osazení

Z grafů je patrné, že nejvíce dostupných pamětí jsou 8 GB moduly, z nichž asi 1/3 je s chlazením. Moduly s chlazením budou primárně použity do sestav určených k zapojení do procesoru. Pro kancelářské sestavy jsou primárně určeny 4 GB moduly, ale těch je momentálně nedostatek. Všechny 4 GB moduly bez chlazení byly již využity při osazování desek, a tak zbylo jen 13 modulů s chlazením. V současnosti je možnost objednat 4 GB moduly, ale bylo rozhodnuto o případném využití 8 GB modulů, nebo neosazení sestavy pamětí RAM. Osazování low-end sestav 8 GB moduly bude probíhat jen ve specifických případech (novější desky). DDR3 paměti budou vzhledem k zastaralosti využity k osazování low-end sestav případně až jako poslední.

Je dostupných celkem 47 (G3900, G3930, G4400) CPU k osazování low-end sestav a 17 (R5 1600, i3-6100) CPU do sestav určených do procesoru. Primárně budou využity procesory AMD R5 1600 a Intel Core i3. Poté se dále budou osazovat sestavy se staršími CPU od Intelu.

POČET KS by SOCKET and TYP DESKY

SOCKET ● LGA1151 ● AM4 ● N/A INTEL

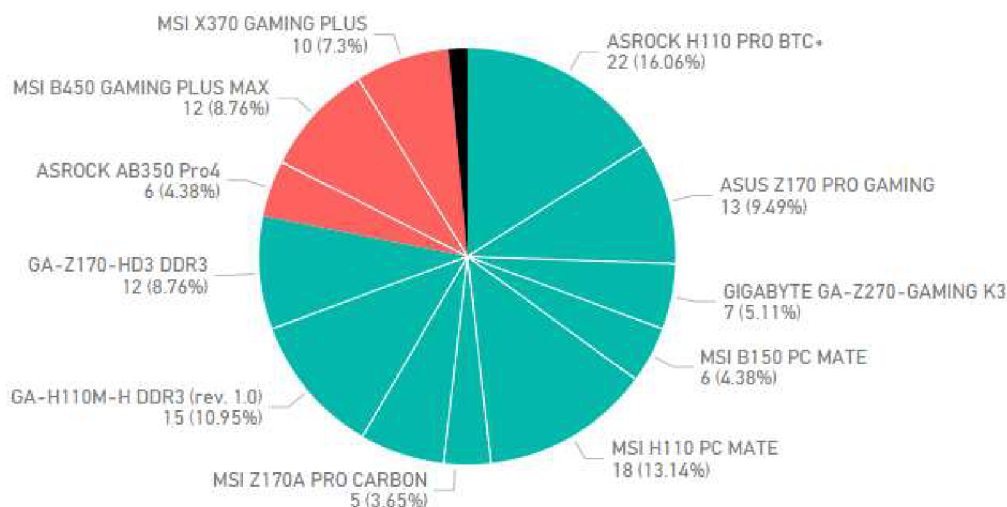


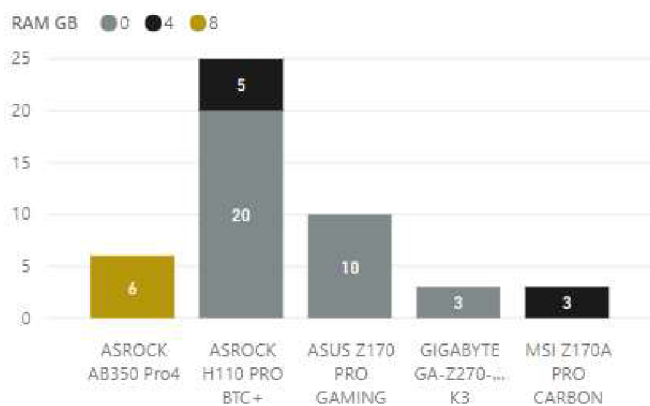
Schéma 14 – Základní desky určené k osazení

Základní desky uvedené na grafu jsou otestované a funkční. Mimo to je ve skladu ještě přibližně stejné množství dalších desek, které budou testovány po zpracování funkčních. Dostupných je celkem 28 desek k osazení procesorem AMD R5 1600 a 107 desek s patičkou LGA1151 pro procesory od Intelu. Také byly objeveny 2 desky s napevno zabudovaným notebookovým procesorem Intel a SODIMM RAM pamětí, tyto desky jsou přímo určeny na těžbu kryptoměn a budou proto odeslány s AMD R5 a Core i3 sestavami.

Kolik sestav již bylo zpracováno a kolik je ještě možné vytvořit?

Vzhledem k tomu, že jsou desky ve skladu delší dobu byly některé již otestovány i osazeny, až nyní ale nastala možnost je zabalit a odeslat na jiná místa. Z grafů je vidět, že bylo osazených 6 desek procesorem AMD R5 1600 a 8 GB RAM modulem. Rovněž bylo celkem 41 desek osazeno procesory Intel G3900, G3930 a G4400, pouze 8 z nich je ale osazeno 4 GB modulem. Protože je dostupných ještě 13 modulů o kapacitě 4 GB s chlazením, bylo rozhodnuto osadit jimi desky ASUS Z170 PRO GAMING a GA-Z270-GAMING K3.

POČET KS by DESKA and RAM GB



POČET KS by TYP CPU

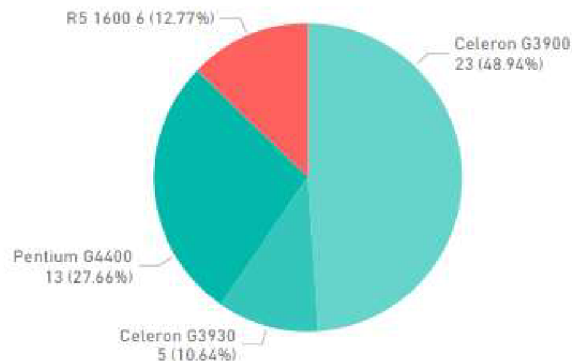
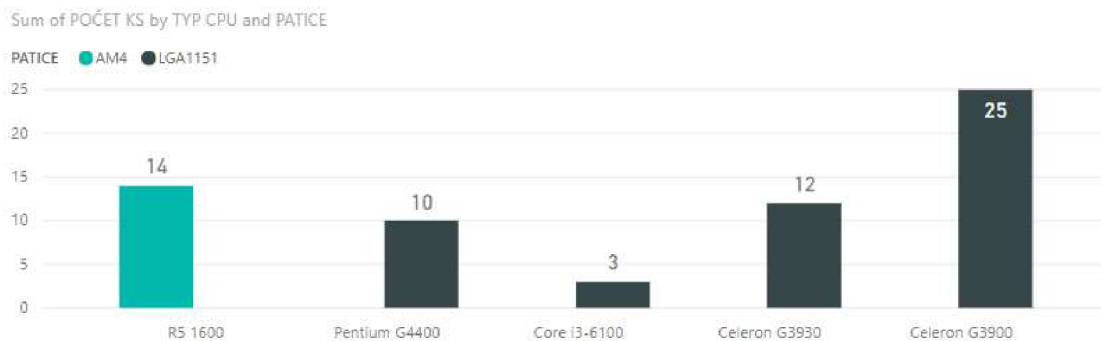


Schéma 15 – Otestované a osazené sestavy

Jako další se budeme věnovat deskám, které teprve budou osazeny CPU a RAM. Porovnávat budeme v tomto případě data z několika tabulek najednou. Z vizualizací a tabulky (byla použita pro lepší přehled parametrů desek) vyplývá, že je možné osadit ještě 14 desek procesory AMD R5. K tomu byly vybráno všech 12 desek MSI B450 GAMING PLUS MAX a 2 z 6 desek ASROCK AB350 Pro4. Ostatní AM4 desky budou zatím uloženy ve skladu a osazeny budou až se doobjednají procesory. Mimo to byly vybrány 3 z 9 desek ASUS B250 MINING EXPERT k osazení procesory Intel Core i3. Zmíněné desky budou osazeny 8 GB RAM modulem. Dále je dostupných 47 procesorů typu Pentium/Celeron a těmi bude osazeno všech 22 desek ASROCK H110 PRO BTC+, všech 7 desek GIGABYTE-GA-Z270-GAMING-K3 (tyto dva typy byly osazovány již dříve) a dále všech 18 desek MSI H110 PC MATE. Desky GA-Z270-GAMING-K3 budou osazeny 8 GB RAM modulem, ostatní budou kompletovány bez RAM.

Celkově můžeme osadit ještě 64 desek (17 z nich určených na těžbu a zbylých 47 určených k prodeji), limitování jsme v tomto případě počtem dostupných procesorů. V plánu je v současnosti objednat více procesorů AMD R5 1600 a Intel Core i3/i5 a pokračovat v osazování desek společně s 8 GB moduly.



POČET KS by TYP MODULU and KAPACITA



TYP DESKY	POČET KS	SOCKET	DDR TYP
ASROCK AB350 Pro4	6	AM4	DDR4
MSI B450 GAMING PLUS MAX	12	AM4	DDR4
MSI X370 GAMING PLUS	10	AM4	DDR4
ASROCK H110 PRO BTC+	22	LGA1151	DDR4
ASUS Z170 PRO GAMING	13	LGA1151	DDR4
GIGABYTE GA-Z270-GAMING K3	7	LGA1151	DDR4
MSI B150 PC MATE	6	LGA1151	DDR4
MSI H110 PC MATE	18	LGA1151	DDR4
MSI Z170A PRO CARBON	5	LGA1151	DDR4
ASUS B250 MINING EXPERT LGA1151 DDR4	9	LGA1151	DDR4
GA-H110M-H DDR3 (rev. 1.0)	15	LGA1151	DDR3
GA-Z170-HD3 DDR3	12	LGA1151	DDR3
ONDA D1800 BTC + Intel CPU napevno + 4GB RAM SODIMM	2	N/A INTEL	SODIMM
Total	137		

Schéma 16 – Výchozí data k další tvorbě sestav

Jaká je situace ohledně nefunkčních a poškozených komponent?

Nyní budou vizualizována data z tabulky komponent, které byly buď poškozeny fyzicky (přeprava, nevhodné skladování, neodborná manipulace) a uvedeny nejčastější druhy poškození a jednotlivé druhy poškozených komponent. V rámci analýzy jsou rozlišovány nefunkční, anebo jinak poškozené komponenty. Nefunkční znamená, že komponenty jsou vizuálně v pořádku ale při zapojení nefungují. Ostatní kategorie zahrnují mechanické poškození pinů na desce, patice a ohořelá a chybějící tranzistory.

Z grafů vyplývá, že v naprosté většině případů jsou poškozené základní desky. To bylo zapříčiněno neodbornou manipulací ze stran brigádníků a nedůsledným balením základních desek určených k přepravě. Za následek to mělo, že značná část základních desek byla poškozena mechanicky, hlavně se jedná o desky s paticemi LGA1151, které jsou náchylné poškození a pokud nejsou osazeny procesorem vyžadují krytku. Dále se jedná o poškození pinů sloužících k připojení periferií na desku. Celkově bylo takto poškozeno 93 základních desek. Ostatní komponenty

byly většinou poškozeny při používání. Celkově se jedná o 22 x GPU, 2 x CPU a 13 x RAM.

Také je patrné, že z velké části jsou zastoupeny komponenty, které jsou vizuálně v pořádku, ale nefunkční. Vzhledem k plánovanému testování komponent je jisté že jejich podíl v rámci celku poroste, zejména se bude jednat o GPU. Momentálně je nutné se nefunkčních komponent zbavit a již je zařízen jejich odvoz. Uvolní se tak místo pro skladování dalších komponent. O tom, zda budou nefunkční komponenty prodány, nebo zlikvidovány bude rozhodnuto po konzultaci s vedením.

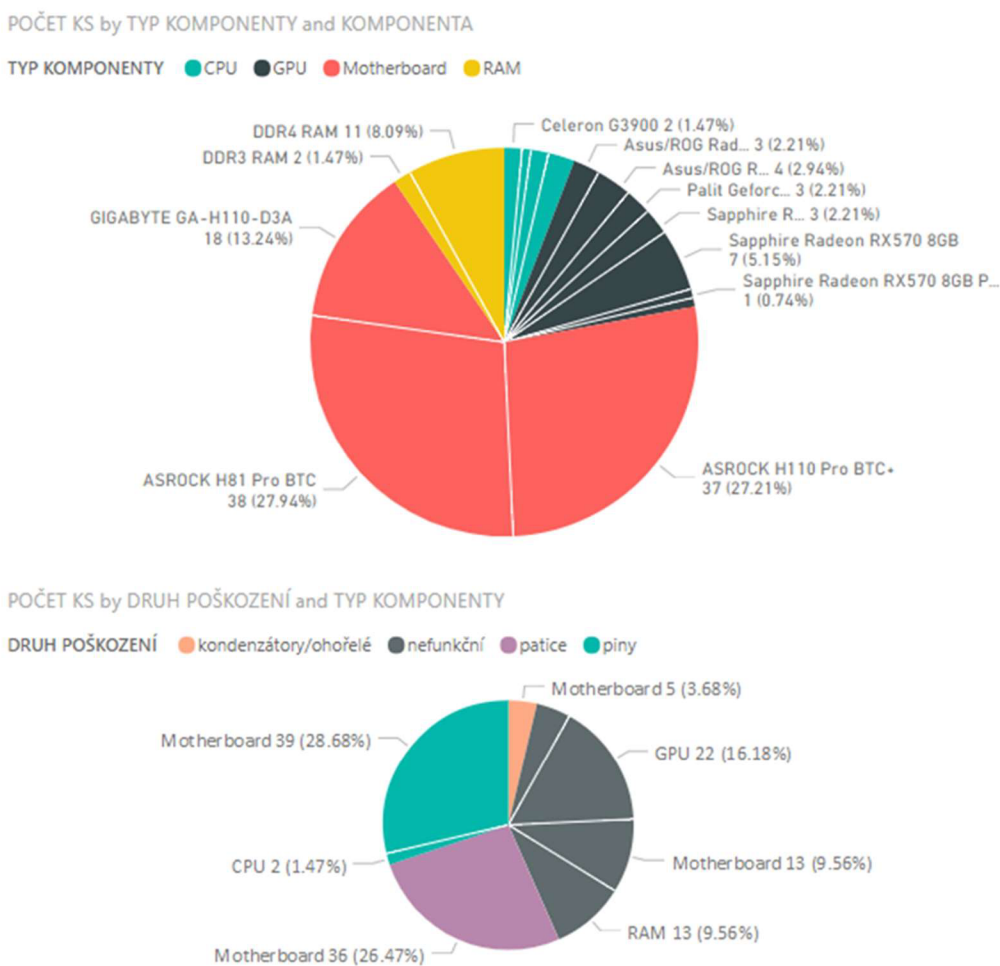


Schéma 17 – Nefunkční a poškozené komponenty

Následují tabulky s výchozími daty:

VÝROBCE KARTY	VÝROBCE ČIPU	GPU ČIP	Velikost paměti
AsRock	AMD Radeon	RX5700	8GB
Asus/ROG	nVidia Geforce	GTX1060	6GB
Asus/ROG	nVidia Geforce	GTX1070	8GB
Asus/ROG	nVidia Geforce	P106	6GB
Asus/ROG	AMD Radeon	R390	8GB
Asus/ROG	AMD Radeon	RX580	4GB
Asus/ROG	AMD Radeon	RX570	4GB
Asus/ROG	nVidia Geforce	GTX1060	6GB
Asus/ROG	AMD Radeon	RX470	4GB
Asus/ROG	AMD Radeon	RX580	8GB
EVGA	nVidia Geforce	GTX1060	6GB
Gigabyte/Aorus	AMD Radeon	RX570	4GB
Gigabyte/Aorus	nVidia Geforce	GTX1060	3GB
Gigabyte/Aorus	nVidia Geforce	GTX1070	8GB
Gigabyte/Aorus	AMD Radeon	R390	8GB
Gigabyte/Aorus	AMD Radeon	RX580	8GB
Inno3D	nVidia Geforce	GTX1080	11GB
Leadtek	nVidia Geforce	GTX1080	11GB
Leadtek	nVidia Geforce	GTX1080	11GB
MSI	nVidia Geforce	GTX1660	6GB
MSI	AMD Radeon	RX470	8GB
MSI	nVidia Geforce	GTX1060	6GB
MSI	nVidia Geforce	GTX1060	6GB
MSI	AMD Radeon	RX480	8GB
MSI	AMD Radeon	RX470	8GB
MSI	nVidia Geforce	GTX1060	6GB
MSI	AMD Radeon	R390	8GB
MSI	nVidia Geforce	GTX1060	6GB
Ostatní/Noname	AMD Radeon	R390	8GB
Palit	nVidia Geforce	P106	6GB
Powercolor	AMD Radeon	RX580	8GB
Powercolor	AMD Radeon	RX570	4GB
Sapphire	AMD Radeon	R370	4GB
Sapphire	AMD Radeon	R380	4GB
Sapphire	AMD Radeon	R380	4GB
Sapphire	AMD Radeon	RX470	4GB
Sapphire	AMD Radeon	RX470	8GB
Sapphire	AMD Radeon	RX480	4GB
Sapphire	AMD Radeon	RX570	4GB
Sapphire	AMD Radeon	RX570	4GB
Sapphire	AMD Radeon	RX570	8GB
Sapphire	AMD Radeon	RX570	8GB
Sapphire	AMD Radeon	RX5700	8GB
Sapphire	AMD Radeon	RX580	4GB
Sapphire	AMD Radeon	RX580	4GB
Sapphire	AMD Radeon	RX580	8GB
Sapphire	AMD Radeon	RX580	8GB
Sapphire	AMD Radeon	VII(Vega20)	16GB
XFX	AMD Radeon	R390	8GB
XFX	AMD Radeon	RX470	8GB
XFX	AMD Radeon	RX480	8GB
XFX	AMD Radeon	RX580	8GB
Zotac	nVidia Geforce	GTX1070	8GB
Zotac	nVidia Geforce	P106	6GB
XFX	AMD Radeon	RX570	8GB
XFX	AMD Radeon	RX580	4GB
Palit	nVidia Geforce	GTX1060	6GB
Zotac	nVidia Geforce	GTX1060	6GB
MSI	nVidia Geforce	GTX1070	8GB
Asus/ROG	AMD Radeon	RX580	8GB
Gigabyte/Aorus	AMD Radeon	RX5700	8GB
AsRock	AMD Radeon	RX5700	8GB
Palit	nVidia Geforce	GTX3060	8GB
Zotac	nVidia Geforce	GTX3060	8GB
XFX	AMD Radeon	RX6700	12GB
Leadtek	nVidia Geforce	GTX1080	11GB

Tabulka 3 – Všechny typy a výrobci karet ve skladu

TYP GPU	PAMĚŤ GB	REFLOW	POČET KS
XFX RX570 8GB 8pin	8	NE	76
XFX RX570 8GB 8pin	8	ANO	13
Zotac GTX1060 6GB AMP!	6	NE	18
MSI GTX1070 8GB ARMOR OC	8	NE	18
ROG RX580 8GB Dual	8	NE	8
Sapphire RX570 8GB	8	NE	8
Gigabyte RX5700 8GB XT	8	NE	5
Palit GTX3060 8GB Ti	8	NE	15
XFX RX470 8GB	8	NE	4
AsRock RX5700 8GB XT Challenger Pro	8	ANO	14
AsRock RX5700 8GB XT Challenger Pro	8	NE	8
Sapphire VII(Vega20) 16GB	16	NE	2
Zotac GTX3060 8GB Ti Twin Edge	8	NE	2
XFX RX6700 12GB XT Speedster quick	12	ANO	1
Palit GTX1060 6GB nebo P106	6	NE	12
Palit GTX1060 6GB nebo P106	6	ANO	4
Sapphire RX580 8GB	8	ANO	8
XFX RX580 8GB	8	ANO	13

Tabulka 4 – Typy GPU z haly 1 v servisu

TYP GPU	REFLOW	PAMĚŤ GB	POČET KS
XFX RX470 8GB	NE	8	131
Sapphire R370 4GB Nitro	NE	4	71
XFX RX580 8GB	NE	8	36
Sapphire R380 4GB Nitro	NE	4	28
Sapphire RX470 8GB	NE	8	12
Asus/ROG RX570 4GB ROG Strix	NE	4	10
Asus/ROG RX580 4GB Dual	NE	4	9
Sapphire RX480 4GB	NE	4	9
XFX RX480 8GB	NE	8	9
MSI GTX1060 6GB Gaming X	NE	6	8
Asus/ROG RX470 4GB ROG Strix	NE	4	6
Sapphire RX580 8GB	NE	8	5
MSI GTX1060 6GB OCV1	NE	6	4
Sapphire RX580 8GB Malé	NE	8	4
Asus/ROG GTX1060 6GB Radiální ventilátor	NE	6	3
Inno3D GTX1080 11GB iChill HerculeZ	NE	11	3
Sapphire RX570 4GB ITX/Compact	NE	4	3
Asus/ROG GTX1060 6GB	NE	6	2
Asus/ROG P106 6GB	NE	6	8
Gigabyte/Aorus RX570 4GB	NE	4	2
Leadtek GTX1080 11GB Radiální ventilátor	NE	11	2
MSI GTX1060 6GB OCV2 (Compact/ITX)	NE	6	1
MSI GTX1660 6GB	NE	6	1
MSI R390 8GB Gaming	NE	8	1
MSI RX470 8GB Gaming X	NE	8	1
MSI RX480 8GB Gaming X	NE	8	1
Powercolor RX570 4GB	NE	4	1
Sapphire R380 4GB ITX/Compact	NE	4	1
Sapphire RX580 4GB	NE	4	1
XFX R390 8GB	NE	8	1
Zotac P106 6GB	NE	6	8

Tabulka 5 – Typy GPU z haly 2 v servisu

TYP MODULU	KAPACITA	POČET KS
DDR3	2GB	8
DDR3	4GB	15
DDR4	4GB	0
DDR4	8GB	44
DDR4 CHLAZENÍ	4GB	13
DDR4 CHLAZENÍ	8GB	20

Tabulka 6 – RAM k osazování sestav

TYP CPU	PATICE	POČET KS
Celeron G3900	LGA1151	25
Celeron G3930	LGA1151	12
Pentium G4400	LGA1151	10
R5 1600	AM4	14
Core i3-6100	LGA1151	3

Tabulka 8 – CPU k osazování sestav

TYP DESKY	SOCKET	DDR TYP	POČET KS
GA-Z170-HD3 DDR3	LGA1151	DDR3	12
GA-H110M-H DDR3 (rev. 1.0)	LGA1151	DDR3	15
ASUS B250 MINING EXPERT LGA1151 DDR4	LGA1151	DDR4	9
ASROCK H110 PRO BTC+	LGA1151	DDR4	22
MSI Z170A PRO CARBON	LGA1151	DDR4	5
MSI B150 PC MATE	LGA1151	DDR4	6
MSI H110 PC MATE	LGA1151	DDR4	18
ASUS Z170 PRO GAMING	LGA1151	DDR4	13
GIGABYTE GA-Z270-GAMING K3	LGA1151	DDR4	7
ONDA D1800 BTC + Intel CPU napevno + 4GB RAM SODIMM	N/A INTEL	SODIMM	2
MSI X370 GAMING PLUS	AM4	DDR4	10
MSI B450 GAMING PLUS MAX	AM4	DDR4	12
ASROCK AB350 Pro4	AM4	DDR4	6

Tabulka 7 – Základní desky pro sestavy

DESKA	CPU	TYP CPU	RAM	RAM GB	POČET KS
ASROCK AB350 Pro4	ANO	R5 1600	ANO	8	6
ASROCK H110 PRO BTC+	ANO	Celeron G3900	ANO	4	5
MSI Z170A PRO CARBON	ANO	Celeron G3900	ANO	4	3
ASUS Z170 PRO GAMING	ANO	Pentium G4400	NE	0	10
GIGABYTE GA-Z270-GAMING K3	ANO	Pentium G4400	NE	0	3
ASROCK H110 PRO BTC+	ANO	Celeron G3900	NE	0	15
ASROCK H110 PRO BTC+	ANO	Celeron G3930	NE	0	5

Tabulka 9 – Kompletní sestavy

KOMPONENTA	TYP KOMPONENTY	DRUH POŠKOZENÍ	POČET KS
GIGABYTE GA-H110-D3A	Motherboard	kondenzátory/ohořelé	4
GIGABYTE GA-H110-D3A	Motherboard	patice	11
ASROCK H81 Pro BTC	Motherboard	patice	16
ASROCK H81 Pro BTC	Motherboard	piny	17
ASROCK H81 Pro BTC	Motherboard	kondenzátory/ohořelé	1
ASROCK H110 Pro BTC+	Motherboard	piny	22
ASROCK H110 Pro BTC+	Motherboard	patice	9
Ryzen 5 1600	CPU	nefunkční	1
Celeron G3900	CPU	nefunkční	2
Celeron G3930	CPU	nefunkční	1
Pentium G4400	CPU	nefunkční	2
Ryzen 5 1600	CPU	piny	2
GIGABYTE GA-H110-D3A	Motherboard	nefunkční	3
ASROCK H110 Pro BTC+	Motherboard	nefunkční	6
ASROCK H81 Pro BTC	Motherboard	nefunkční	4
DDR3 RAM	RAM	nefunkční	2
DDR4 RAM	RAM	nefunkční	11
XFx Radeon RX580 8GB	GPU	nefunkční	1
Sapphire Radeon RX480 4GB	GPU	nefunkční	3
Sapphire Radeon RX570 8GB	GPU	nefunkční	7
Asus/ROG Radeon RX580 8GB Dual	GPU	nefunkční	3
Asus/ROG Radeon RX580 8GB Dual PO REFLOW	GPU	nefunkční	4
Sapphire Radeon RX570 8GB PO REFLOW	GPU	nefunkční	1
Palit Geforce GTX1060 6GB PO REFLOW	GPU	nefunkční	3

Tabulka 10 – Záznamy o poškozených a nefunkčních komponentách

7 Shrnutí výsledků

Během zpracování práce byla zpřehledněna situace ve skladovém hospodářství Hradecké pobočky a díky tomu bylo rozhodnuto o krocích vedoucích k zefektivnění práce a vyprázdnění skladu pobočky v HK.

Konkrétně bylo zjištěno, kolik je možné vytvořit PC sestav s různými parametry a díky tomu je rychleji zkompletovat a odeslat je do procesu nebo k prodeji. Byla zpřehledněna situace týkající se grafických karet a byl jasně definován postup práce s nimi v rámci servisu. Také je jasně dané pořadí, v jakém se budou jednotlivé GPU testovat a sestavy osazovat. Díky těmto informacím bude několik následujících týdnů či měsíců (v závislosti na docházce brigádníků) probíhat práce efektivněji.

Bylo zjištěno, jaké typy komponent jsou nedostatkové, a proto bude po domluvě s vedením rozhodnuto o jejich případném objednání.

Také byly zjištěny nejčastější typy poškození komponent a na základě toho bude vytvořen školící dokument pro zaměstnance, aby se předcházelo ničení komponent.

V rámci kontroly financí bylo zjištěno, že se daří dodržovat stanovený měsíční limit a také jsou jasně definovány položky, za které se nejvíc utrácí. Na základě analýzy dat z tabulky financí budou zavedeny jiné postupy při zapisování nakupovaného zboží. Z dat v tabulce zaměstnanců byly díky vizualizacím odvozeny informace, které by v případě potřeby pomohly s logistikou komponent a plánováním práce.

8 Závěry a doporučení

V práci byla představena oblast vizualizace dat a byly uvedeny nejdůležitější pojmy související s touto oblastí. Vzhledem k zachování konzistence byly nejprve představeny teoretické pojmy a důvody, proč je vhodné data vizualizovat. Také byl uvedený historický význam a kontext vizualizace dat a představena vizualizace dat za využití rozšířené reality.

Zásady pro tvorbu vizualizací a nebezpečí zavádějících vizualizací a statistik byly představeny v další části, stejně jako nejrůznější typy grafů, diagramů a schémat. Dále byly představeny různé typy programů a jiné technologie pro vizualizaci dat. Vybrané prezentované programy byly v praktické části doplněny o aplikaci ve firemním prostředí.

Mohly být představeny i další technologie vizualizace, například Tableau Desktop, nebo by se práce mohla podrobněji věnovat popsaným technologiím (AnyChart, JazykR). Při výběru technologií pro zpracování praktické části byla zohledněna zejména dostupnost aplikací (všechny použité technologie byly získány zdarma). Pro potřeby zpracování výsledků byly vybrány vhodné technologie a výsledné vizualizace jasně evokují informace, které jsou klíčové pro koordinaci dalších pracovních postupů.

V budoucnu bude nutné nadále sledovat situaci ohledně komponent, financí a zaměstnanců ve firmě, a tak budou postupy a výsledky uvedené v této práci přínosné při další analýze dat.

9 Seznam použité literatury

- [1] VAUGHAN, Jack. What is data? Tech Target [online]. 2019 [cit. 2022-08-08]. Dostupné z: <https://www.techtarget.com/searchdatamanagement/definition/data>
- [2] KITCHIN, Rob. The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences. Knihy Google [online]. 2014 [cit. 2022-08-08]. Dostupné z: [https://books.google.cz/books?hl=cs&lr=&id=GfOICwAAQBAJ&oi=fnd&pg=PP1&dq=Kitchin,+Rob+\(2014\).+The+Data+Revolution&ots=pewbLYUmTU&sig=Cbu sqAiCdB05k8ZBTctUOHu0eXk&redir_esc=y#v=onepage&q=Kitchin%2C%20Rob%20\(2014\).%20The%20Data%20Revolution&f=false](https://books.google.cz/books?hl=cs&lr=&id=GfOICwAAQBAJ&oi=fnd&pg=PP1&dq=Kitchin,+Rob+(2014).+The+Data+Revolution&ots=pewbLYUmTU&sig=Cbu sqAiCdB05k8ZBTctUOHu0eXk&redir_esc=y#v=onepage&q=Kitchin%2C%20Rob%20(2014).%20The%20Data%20Revolution&f=false)
- [3] GEMIGNANI, Zach, Chris GEMIGNANI, Richard GALENTINO a Patrick Jude SCHUERMANN. Efektivní analýza a využití dat. Přeložil Jiří HUF. Brno: Computer Press, 2015. ISBN 978-80-251-4571-5.
- [4] MYSLIVEC, Jaroslav. Vizualizace vícerozměrných dat. Pardubice: Univerzita Pardubice, Fakulta filozofická, 2012. ISBN 978-80-7395-445-1.
- [5] Typy dat. Wikiskripta [online]. 2018 [cit. 2022-09-08]. Dostupné z: https://www.wikiskripta.eu/w/Typy_dat
- [6] SOCHOR, Jiří; BENEŠ, Bedřich; FELKEL, Petr; ŽÁRA, Jiří. Vizualizace (Visualization). 1. vyd. Praha: FEL ČVUT Praha, 1997. 197 str.
- [7] WARD, Matthew, Georges GRINSTEIN a Daniel KEIM. Interactive Data Visualization: Foundations, Techniques, and Applications. Taylor & Francis Group [online]. 2010 [cit. 2022-09-08]. Dostupné z: <https://www.taylorfrancis.com/books/mono/10.1201/9780429108433/interactive-data-visualization-matthew-ward-georges-grinstein-daniel-keim>
- [8] 4 DŮVODY, PROČ SE JIŽ NEOBEJDETE BEZ VIZUALIZACE DAT. SAS [online]. [cit. 2022-09-08]. Dostupné z: https://www.sas.com/content/dam/SAS/cs_cz/doc/other1/SAS-ebook.pdf
- [9] Tematická mapa. Wikipedia [online]. [cit. 2021-8-19]. Dostupné z: https://cs.wikipedia.org/wiki/Tematick%C3%A1_mapa
- [10] Key Figures in the History of Data Visualization. Medium [online]. 2016 [cit. 2022-08-12]. Dostupné z: <https://medium.com/@Infogram/key-figures-in-the-history-of-data-visualization-30486681844c>
- [11] WENKAI, Li, S. K. ONG a A. Y. C. NEE. A State-of-the-Art Review of Augmented Reality in Engineering Analysis and Simulation. MTI [online]. 2017 [cit. 2022-10-10]. Dostupné z: <https://www.mdpi.com/2414-4088/1/3/17/htm>
- [12] Rozšířená realita ve ŠKODA AUTO zjednodušuje údržbu. Vše o průmyslu [online]. 2021 [cit. 2022-10-10]. Dostupné z:

- <https://www.vseoprmyslu.cz/digitalizace/virtualni-rozsirena-realita/rozsirena-realita-ve-skoda-auto-zjednodusuje-udrzbu.html>
- [13] Proč a jak správně vizualizovat data. *ABRA* [online]. [cit. 2021-8-19]. Dostupné z: <https://www.abra.eu/proc-a-jak-spravne-vizualizovat-data/>
- [14] 3. přednáška: Základní principy vizualizace. MUNI [online]. [cit. 2021-8-20]. Dostupné z: https://is.muni.cz/el/1433/jaro2016/PV251/um/61907197/3_zakladni_principy_vizu_alizace.pdf
- [15] Misleading Graphs: Real Life Examples: Misleading Graphs in Real Life: Overview. Statistics How To [online]. 2022 [cit. 2022-10-08]. Dostupné z: <https://www.statisticshowto.com/probability-and-statistics/descriptive-statistics/misleading-graphs/>
- [16] Alena Schillerová – Příspěvky. Facebook [online]. 2019, 14-02-2019 [cit. 2022-08-15]. Dostupné z: <https://www.facebook.com/SchillerovaAlena/posts/596697190803564>
- [17] CEMPER, Jan. PRO POBAVENÍ: Grafy Andreje Babiše. Manipulátoři.cz [online]. 2019, 15.12.2019 [cit. 2022-08-29]. Dostupné z: <https://manipulatori.cz/pro-pobaveni-grafy-andreje-babise/>
- [18] What are graphs?. Quora [online]. [cit. 2022-08-08]. Dostupné z: <https://www.quora.com/What-are-graphs-1>
- [19] What is the difference between chart and graph?. Quora [online]. [cit. 2022-08-08]. Dostupné z: <https://www.quora.com/What-is-the-difference-between-chart-and-graph>
- [20] Data Viz Project. Data Viz Project [online]. [cit. 2022-08-08]. Dostupné z: <https://datavizproject.com/>
- [21] Svaz průmyslu a dopravy ČR na Twitteru: „České zboží si v roce 2018 mohli koupit lidé ve více než 200 zemích světa. Exportéři vyvezli zboží a služby za skoro 4,4 mld. Kč. Twitter [online]. 2019 [cit. 2022-08-08]. Dostupné z: https://pbs.twimg.com/media/DyzSi_FXcAArZGL.jpg
- [22] Contingency table. Wikipedia [online]. 2021 [cit. 2022-09-08]. Dostupné z: https://en.wikipedia.org/wiki/Contingency_table
- [23] Microsoft Excel. Wikipedia [online]. 2022 [cit. 2022-10-08]. Dostupné z: https://cs.wikipedia.org/wiki/Microsoft_Excel
- [24] What is Power BI?. Microsoft Docs [online]. 2022 [cit. 2022-08-12]. Dostupné z: <https://docs.microsoft.com/en-us/power-bi/fundamentals/power-bi-overview>
- [25] Microsoft Power BI. Wikipedia [online]. 2022 [cit. 2022-11-08]. Dostupné z: https://en.wikipedia.org/wiki/Microsoft_Power_BI

- [26] SNIPES, Genifer. Product Review Google Data Studio [online]. [cit. 2022-08-12]. Dostupné z: https://www.researchgate.net/publication/323325377_Product_Review_Google_Data_Studio
- [27] JavaScript Charting Library - AnyChart - JS charts. Anychart [online]. 2021 [cit. 2021-08]. Dostupné z: <https://www.anychart.com/products/anychart/overview/>
- [28] Walrus - Graph Visualization Tool. CAIDA [online]. 2021 [cit. 2021-08]. Dostupné z: <https://www.caida.org/catalog/software/walrus/>
- [29] R (programovací jazyk). *Wikipedia* [online]. [cit. 2021-8-30]. Dostupné z: [https://cs.wikipedia.org/wiki/R_\(programovac%C3%AD_jazyk\)](https://cs.wikipedia.org/wiki/R_(programovac%C3%AD_jazyk))
- [30] Data visualization with R and ggplot2. *R Graph Gallery* [online]. [cit. 2021-8-30]. Dostupné z: <https://www.r-graph-gallery.com/ggplot2-package.html>

10 Přílohy

Příloha 1 – Zadání bakalářské práce

Zadání bakalářské práce

Autor:	Daniel Ott
Studium:	11800443
Studijní program:	B6209 Systémové inženýrství a informatika
Studijní obor:	Informační management
Název bakalářské práce:	Nástroje vizualizace dat pro podporu rozhodování
Název bakalářské práce A):	Data visualization tools for decision support

Cíl, metody, literatura, předpoklady:

Popsat různé aspekty vizualizace dat se zaměřením na grafy, založené na statistických charakteristikách. Zaměří se na zvolené typy grafů a dat. Budou popsány a prezentovány vybrané technologie.

V praktické části se student bude věnovat vlastním návrhům grafů s vypovídací hodnotou.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis* (2nd ed.). New York, NY: Springer-Verlag. Retrieved from <https://ggplot2.tidyverse.org>

Wilkinson, L. (2005). *The grammar of graphics* (2nd ed.). New York, NY: Springer.

Zinovyev A.: Data visualization in political and social sciences.
<https://arxiv.org/ftp/arxiv/papers/1008/1008.1188.pdf>

<https://theses.cz/id/e9v0gf/STAG85898.pdf>

Data visualization with R and ggplot2. *R Graph Gallery* [online]. <https://www.r-graph-gallery.com/ggplot2-package.html>

Zadávací pracoviště: Katedra informatiky a kvantitativních metod,
Fakulta informatiky a managementu

Vedoucí práce: prof. RNDr. Hana Skalská, CSc.

Oponent: Ing. Kateřina Frončková, Ph.D.

Datum zadání závěrečné práce: 26.1.2021