

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



ÚVOD DO VIZUALIZÁCIE INFORMÁCIÍ PRE MANAŽÉROV

Diplomová práca

Bratislava 2019

Veronika Zajceva

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

ÚVOD DO VIZUALIZÁCIE INFORMÁCIÍ PRE MANAŽÉROV

Diplomová práca

Študijný program: Manažérska matematika
Študijný odbor: 1114 Aplikovaná matematika
Školiace pracovisko: Katedra algebry a geometrie
Vedúci práce: doc. RNDr. Andrej Ferko, PhD.

Bratislava 2019

Veronika Zajceva



86603461

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

- Meno a priezvisko študenta:** Bc. Veronika Zajceva
Študijný program: manažérska matematika (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: aplikovaná matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický
- Názov:** Úvod do vizualizácie informácií pre manažérov
Introduction to information visualization for managers
- Anotácia:** Vysvetlíme históriu a základy vizualizácie dát a procesov. Zvolíme niekoľko motivačných príkladov. Navrhujeme vizualizačnú prezentáciu na vybraných dátach o kryptomene Bitcoin. Na implementáciu zvolíme vhodný autorský nástroj (napr. R, Tableau).
- Cieľ:** Vytvoriť výučbový materiál o vizualizácii informácií pre predmet Počítačová grafika pre manažérov.
- Literatúra:** Domik, G. 1999. Tutorial on Visualization. [online] <http://www6.uniovi.es/hypvis/domik/folien.html>.
Chi, EH. 2000. A Taxonomy of Visualization Techniques Using the Data State Reference Model. [online] <http://www.ics.uci.edu/~kobsa/courses/ICS280/InfoViz2000/ed-chi.pdf>.
Tufte, ER. 2001. The Visual Display of Quantitative Information, Cheshire: Graphics Press.
Inselberg, A. 2009. Parallel Coordinates - Visual Multidimensional Geometry and Its Applications, New York: Springer.
Vaganyan, GA. 1985. Mašinnaja grafika v upravľenii. Jerevan: Ajastan.
- Kľúčové slová:** vizualizácia informácií, metóda paralelných súradníc, vizualizácia viacrozmerných dát
- Vedúci:** doc. RNDr. Andrej Ferko, PhD.
Katedra: FMFI.KAG - Katedra algebry a geometrie
Vedúci katedry: doc. RNDr. Pavel Chalmovianský, PhD.
- Dátum zadania:** 09.11.2017
- Dátum schválenia:** 04.12.2017
- prof. RNDr. Michal Fečkan, DrSc.
garant študijného programu

.....
študent.....
vedúci práce



Pod'akovanie

Touto cestou by som sa chcela poďakovať svojmu vedúcemu práce, doc. RNDr. Andrejovi Ferkovi, PhD., za vytvorenie priestoru na spracovanie zaujímavej témy. Ďakujem za cenné rady a usmerňovanie pri písaní diplomovej práce, za všetok čas strávený pri konzultáciách a za ochotu odpovedať na moje otázky. Chcela by som sa poďakovať aj spoločnosti Tableau za možnosť využívať ich vizualizačný softvér na tvorbu vizualizácií a ich následnú analýzu. Zároveň moje poďakovanie patrí aj spoločnostiam CoinMarketCap a Blockchain, vďaka ktorým som mohla využiť dáta o kryptomene Bitcoin, za účelom uskutočnenia praktickej časti v diplomovej práci.

Abstrakt v štátnom jazyku

ZAJCEVA, Veronika: Úvod do vizualizácie informácií pre manažérov [Diplomová práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra algebry a geometrie; školiteľ: doc. RNDr. Andrej Ferko, PhD., Bratislava, 2019, 99 s.

V diplomovej práci sa zaoberáme tvorbou vizualizácií informácií a ich využitím nielen v oblasti manažmentu. Predstavujeme rôzne druhy grafických prezentácií dát z osemdesiatych rokov, ku ktorým ponúkame na výber novodobé ekvivalenty. Opisujeme, ako aplikovať jednotlivé druhy do praxe na základe povahy zobrazovanej informácie. S pomocou dát o kryptomene Bitcoin, získaných na webstránkach spoločností CoinMarketCap a Blockchain, vytvárame vizualizačné príklady dvoj-, troj-, štvor- a viacrozmerných dát. Na tento účel používame softvéry R a Tableau, a taktiež čerpáme poznatky zo štatistiky či časových radov. Cieľom práce je priblížiť čitateľovi tému vizualizácií informácií, a navrhnúť grafickú reprezentáciu viacrozmerných dát metódou paralelných súradníc. Na zvýšenie informačnej hodnoty práce prepájame vysvetlenú teóriu s praktickou časťou.

Kľúčové slová: vizualizácia informácií, metóda paralelných súradníc, vizualizácia viacrozmerných dát

Abstract

ZAJCEVA, Veronika: Introduction to information visualization for managers [Diploma Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Algebra and Geometry; supervisor: doc. RNDr. Andrej Ferko, PhD., Bratislava, 2019, 99 p.

In this diploma thesis we deal with the creation of information visualization and its usage not only in the field of management. Introducing various types of graphical presentations of data from the eighties, to which we offer a choice of modern equivalents. We describe how to apply each type to practice based on the information properties displayed. Using data about cryptocurrency Bitcoin, gained from websites CoinMarketCap and Blockchain, we create visualization examples of two-, three-, four- and multidimensional data. For this purpose, we use R and Tableau software, as well as learning from statistics and time series. The aim of the thesis is to introduce the reader to the topic of information visualization and to design a graphical representation of multidimensional data using method of parallel coordinates. To increase the information value of the work, we link the explained theory to the practical part.

Keywords: information visualization, parallel coordinates, multidimensional data visualization

Obsah

Úvod.....	9
1 O čom hovoríme, keď hovoríme o vizualizácii	11
1.1 Z histórie.....	12
1.2 Druhy vizualizácií.....	14
1.2.1 Ilustračný graf charakteristík objektu	15
1.2.2 Ilustračný graf funkčných závislostí.....	15
1.2.3 Porovnávacie diagramy	16
1.2.4 Dynamické diagramy.....	17
1.2.5 Štatistické mapy	18
1.2.6 Plánový graf.....	19
1.2.7 Sieťové grafy.....	19
1.3 Referenčný model stavu informácií	20
1.4 Paralelné súradnice.....	21
1.5 Otázky vizuálneho myslenia	24
1.5.1 Štatistické pojmy	26
1.5.2 Časové rady	28
2 Špecifikácia praktickej časti.....	29
2.1 Dáta o Bitcoine	29
2.2 Vizualizačný nástroj Tableau	31
2.2.1 Získanie študentskej licencie	31
2.2.2 Prvý kontakt s Tableau Desktop.....	32
2.3 Štatistický nástroj R.....	34
3 Vizualizácie v praxi	35
3.1 Čiastočná vizualizácia viacrozmerných dát	35
3.1.1 Dvojrozmerné vizualizácie.....	35
3.1.2 Trojrozmerné vizualizácie	42
3.1.3 Štvorrozmerné vizualizácie	43
3.2 Tvorba paralelnej súradnicovej sústavy	45
4 Vyhodnotenie praktickej časti.....	52
4.1 Interpretácia vizualizácie metódou paralelných súradníc	52

4.2 Aplikovanie referenčného modelu stavu informácií.....	54
Záver.....	57
Zoznam použitej literatúry	59
Príloha A.....	61
Novodobé ekvivalenty k vizualizáciám informácií	61
Príloha B	84
Obrázkový postup tvorby paralelnej súradnicovej sústavy	84

Úvod

Žijeme v období, kedy máme jednoduchý prístup k veľkému množstvu rozličných dát. Vďaka technologickému pokroku dokážeme získavať informácie v ľubovoľnej oblasti záujmu. V súčasnosti už nepredstavuje zásadný problém otázka, kde a akým spôsobom získavať dáta. Do popredia sa dostáva aktuálnejšia úloha – ako vyťažiť z dostupných dát čo najviac užitočných informácií.

Manažéri v podnikoch musia na dennej báze robiť rozhodnutia, s cieľom zlepšiť chod firmy. Avšak, častokrát nemajú dostatok času na skúmanie každého problému do hĺbky. V takom prípade je pre nich vhodné vizualizovať dostupné informácie. Manažéri sa dokážu relatívne rýchlo zorientovať v problematike, porozumieť základným súvislostiam a vyvodit' správne rozhodnutie na základe prehľadných grafov s vysokou výpovednou hodnotou. Z tohto praktického dôvodu sme sa rozhodli v diplomovej práci venovať téme vizualizácie informácií. Inšpiráciou na spracovanie tejto témy bolo aj dielo [1], v ktorom Vaganyan uviedol rôzne druhy grafov používaných v manažmente. V práci predstavíme jednotlivé typy vizualizácií, ponúkneme k nim novodobé ekvivalenty a načrtneme, ako z nich efektívne získavať informácie. Taktiež sa zaoberáme problémom vizualizácie viacrozmerných dát. Jedno z komplexnejších riešení vychádza z práce Inselberga o metóde paralelných súradníc [2]. Pomocou tejto metódy vytvoríme vizuálnu reprezentáciu dát z finančnej oblasti vo vybranom vizualizačnom nástroji.

Cieľom práce je vytvoriť pre čitateľa vzdelávací materiál, v ktorom nájde základný prehľad na tému vizualizácie informácií. Táto práca má za úlohu rozšíriť obzory o histórii vizualizácií, jednotlivých typoch grafov a vizualizačných metódach. Na upevnenie teoretických znalostí ponúkneme konkrétne príklady vizualizácií a uvedieme teóriu do praxe. V praktickej časti používame na tvorbu vizualizácií dáta o kryptomene Bitcoin, získané z webových stránok spoločností CoinMarketCap [3] a Blockchain [4]. Z vizualizačných softvérov budeme vo veľkej miere využívať Tableau [5] a štatistický nástroj R [6]. Časťou cieľa sa javí aj vytvorenie vizualizácie viacrozmerných dát metódou paralelných súradníc, keďže táto funkcionálna chyba v predvolenom balíku možností Tableau.

Pri vytváraní grafov budeme brať do úvahy aj charakter zhromaždených dát. Tvorbou vizualizácií chceme pomôcť čitateľovi v hľadaní odpovedí na niektoré otázky. V prípade dát o Bitcoine môžeme z vytvorených obrazov napríklad zisťovať, či klesajúca trhová cena má

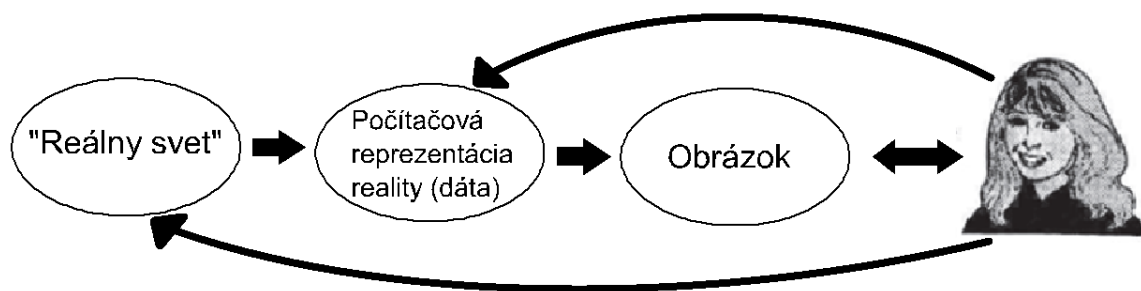
vplyv na zvýšenie dominancie Bitcoinu spomedzi ostatných kryptomien. Istotne si spomíname aj na extrémny nárast trhovej ceny Bitcoinu na konci roku 2017. Medzi hypotetické otázky zaradíme aj také, ktoré pojednávajú o tom, či bolo možné predvídať tak výrazné kolísanie cien, prípadne či sa môže podobná situácia v budúcnosti zopakovať. Dôležitú súčasť vykonanej práce tvorí analýza vizualizácií informácií, preto jej venujeme v práci značnú pozornosť.

Diplomová práca je členená do štyroch kapitol. Prvá kapitola predstavuje úvod do problematiky. Čitateľovi v nej priblížime históriu vizualizácií, predstavíme základné druhy vizualizácií a vysvetlíme dôležité pojmy, ktoré úzko súvisia s danou témou. V druhej kapitole špecifikujeme reálne dáta a vizualizačné nástroje, s ktorými budeme pracovať v praktickej časti diplomovej práce. V tretej kapitole venujeme pozornosť vizualizáciám v praxi a budeme využívať vybrané dáta a softvéry na tvorbu konkrétnych grafov. Zároveň, poskytneme kompletný návod na vytvorenie viacrozmernej vizualizácie metódou paralelných súradníc v konkrétnom vizualizačnom softvéri. Štvrtá kapitola je určená na vyhodnotenie praktickej časti a doplnenie informácií k téme vizualizácií, ktoré vyplynuli z obsahu diplomovej práce.

1 O čom hovoríme, keď hovoríme o vizualizácii

Ľudia zaznamenávajú realitu v podobe obrázkov už od dávnej minulosti. Svedčia o tom napríklad jaskynné maľby, babylonské mapy zobrazené na hlinených tabuliach či ilustrácie v zápiskoch Leonarda da Vinciho. Pomocou obrázkov sa ľudia snažia lepšie porozumieť reálnemu svetu a rýchlejšie si dať informácie do súvislostí.

Vizualizácia je metóda výpočtu. Premieňa symbolický zápis na geometrický, čím nám umožňuje skúmať dáta, simulácie a výpočty. Vďaka vizualizácii dokážeme získať nadhľad nad problémom či zbadáť riešenie, ktoré sme pred vizualizáciou dát neboli schopní vidieť. Odhaduje sa, že 50% mozgových neurónov je spojených práve s videním, a teda cieľom vizualizácie vo vedeckých výpočtoch je aktivovanie týchto neurónov [7]. Prostredníctvom vizuálneho vnímania nadobúdame informácie, ktoré sa na zlomky sekúnd ukladajú v sensorickej pamäti. Tieto informácie pripomínajú fotografie, preto aj vo vytváraní obrázkov metódami počítačovej grafiky dokážeme nájsť analógiu s fotografovaním [8]. Pomocou vizualizácie vytvárame grafický obsah, v ktorom dokážeme komunikovať určitú informáciu. Je to proces transformácie informácií do vizuálnej podoby. V súčasnosti má vizualizácia široké využitie vo vedecko-technickej oblasti, vo vzdelávaní, v počítačovej grafike a v mnohých ďalších obchodných aj odborných oblastiach.



Obr. 1 - Grafické vysvetlenie vzťahov v procese vizualizácie [9]

Proces vizualizácie podľa [9] je mapovanie počítačovej reprezentácie na vnemovú – vizuálnu – reprezentáciu. Realitu môžeme sledovať pomocou dát – napr. úhrn zrážok v priebehu dňa, vývoj cien Bitcoinu, dosah reklamnej kampane na sociálnej sieti Facebook. Namerané dáta následne premieňame do obrázkov v procese zobrazenia, aby sme maximalizovali ľudské porozumenie a komunikáciu reality. Vizuálnym vnímaním pozorovateľ dokáže z obrázka získať obrazový aj obsahový druh informácie. Vizualizácia nám umožňuje skúmať namerané

dáta, odhaľovať skryté súvislosti, lepšie chápať procesy, získať nové pohľady, kontrolovať kvalitu vykonaných simulácií a meraní, a mnoho ďalšieho [9]. Vzťah pozorovateľa a reality, resp. jej reprezentácie je zobrazený na Obr. 1.

1.1 Z histórie

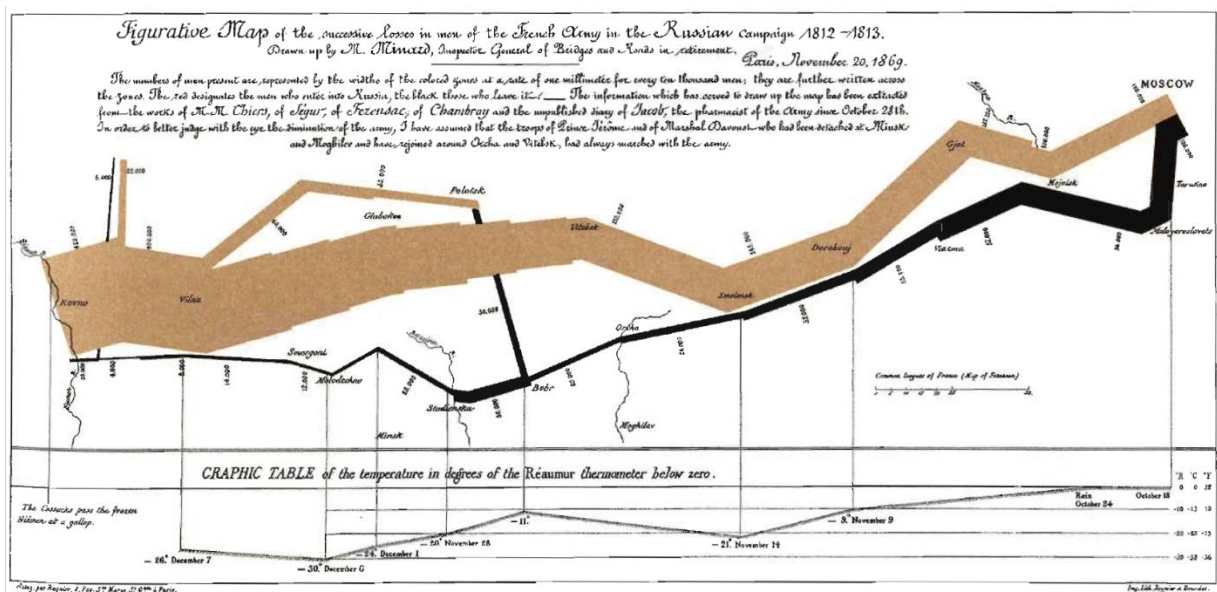
Donedávna sa realita zaznamenávala iba v číselnej podobe a pomocou tabuliek. Grafická reprezentácia dát začala vznikáť dlho po tom, čo boli objavené logaritmy, karteziánske súradnice, derivácie, limity či teória pravdepodobnosti. Prvé základy vizualizácie dát a informácií sa začali klásť v druhej polovici 18. storočia. Skúmalo sa, ako dodať nameraným dátam a z nich získavaným informáciám abstraktnejšiu podobu, ktorá by ich dokázala jednoducho prezentovať a vysvetliť pozorovateľovi. Spojením umeleckých, štatistických a matematických schopností začali dáta postupne nadobúdať grafickú podobu, ako ju poznáme v súčasnosti. Jednotlivé čísla v dátových súboroch, ako napríklad množstvo, časový rad, rozptyl, sa začali znázorňovať ako body, priamky, krivky, symboly, v súradnicovej sústave či farebne sa rozlišovať [10].

Tufte vo svojej publikácii [10] navrhuje niekoľko odporúčaní, ktoré udržia štatistické vizualizácie prehľadnými. Odporúča zobrazovať také dáta, ktoré podnecujú pozorovateľa navzájom ich porovnávať a skúmať. Musí sa však dodržať zmysluplnosť zobrazovaných informácií, aby výsledné vizualizácie prinášali ľuďom pridanú hodnotu. Vizualne príťažlivý graf nemá zmysel, ak sú v ňom zobrazené spolu nesúvisiace dáta, napríklad vývoj cien akcií na londýnskej burze a priemerná denná teplota vzduchu v Paríži v rovnakom období. Vizualizácia by mala postupne odhaľovať informácie od celkového pohľadu až k najmenším detailom, pričom by mal byť presne stanovený cieľ, ktorý má plniť. Vizualizácie môžu slúžiť na opísanie a pochopenie dát, na ich skúmanie či na prípravu podkladov na ďalšiu prácu s týmito dátami.

Publikácia [10] obsahuje zbierku rozličných druhov vizualizácií dát a informácií. Tufte v nej, okrem iného, uvádza pozoruhodné zobrazenia prevažne štatistickej informácie, ktoré vznikli pred desiatkami či stovkami rokov. Za základ, na ktorej stavia teóriu informačnej grafiky, berie dátové mapy, zobrazenia časových radov, grafy zobrazujúce časopriestor, či relačné grafy.

Snáď jeden z najlepšie nakreslených štatistických grafov od vzniku vizualizácie informácií, je mapa zobrazujúca osudovú výpravu Napoleonových vojsk do Ruska, Obr. 2. Charles Joseph Minard ju vytvoril v roku 1869 skombinovaním dátovej mapy s časovými radmi. Na geografickej mape zobrazil svetlohnedou krivkou pochod Napoleonovej armády smerom do Moskvy, jej šírka indikuje veľkosť armády počas výpravy. Čiernou krivkou zobrazil návratovú cestu z Moskvy a znižujúcu sa veľkosť armády. Táto návratová krivka je prepojená s grafom teploty vzduchu v prislúchajúcom dni, zobrazeným v spodnej časti vizualizácie.

Obdivuhodnosť tejto mapy spočíva v tom, že Minard dokázal prehľadne zobraziť až šesť rozmerov z dát o Napoleonovej armáde. Geografická informácia je vyjadrená v dvojrozmernom priestore, na ktorom je znázornená presná cesta armády. Dvojfarebná litografia zabezpečila označenie smeru výpravy a šírka kriviek hovorí o veľkosti armády počas výpravy. Zároveň sa mu podarilo evidovať aktuálnu teplotu a k nej prislúchajúci dátum počas návratu armády z Moskvy [10].



Obr. 2 - Šesť rozmerov zobrazených na Minardovej mape Napoleonovej vojenskej výpravy [10]

V súčasnej dobe pribúdajú s rozvojom technológií aj nové zdroje, ktoré dokážu vyprodukovať veľké množstvo informácií. Obrovský objem dát produkujú napríklad superpočítače, vesmírne družice, magnetická rezonancia, atď. Ľudský mozog nedokáže spracovať toľko informácií v číselnej podobe naraz, a teda väčšina získavaných hodnôt sa iba zbiera a ukladá na dátové úložiská. Koncom minulého storočia začala vznikať potreba

interpretovať vhodným spôsobom nevyužitú informáciu. Možnosť, ako sa vysporiadať s príliš veľkým množstvom dát, sa našla vo využití obrázkov. Vzniklo napríklad odvetvie vizualizácie na vedecké účely, tzv. ViSC (z angl. visualization in scientific computing). Vďaka vedecko-technickej vizualizácii si vedci dokážu predstaviť rozsiahle výpočty a simulácie, môžu prichádzať na nové myšlienky a objavy, a dokážu komunikovať o získaných výsledkoch s inými vedcami [7].

Vzniklo taktiež aj odvetvie vizualizácie informácií, tzv. InfoVis (z angl. information visualization). Avšak, jeho rozvoj je o niečo pomalší ako rozvoj ViSC. Dôvodom je hlavne povaha dát. Zatiaľ čo vo vedeckej oblasti sú vizualizácie využívané najmä na zobrazenie dát s fyzikálnym alebo geometrickým základom, ako napríklad elektromagnetické pole, tok tekutín či geografia, tak v oblastiach, kde nemáme dost' presné matematické modely, chýba dátam takáto štruktúra. V odboroch ako riadenie procesov, komunikácie, financie a podobne, sa vyskytuje veľa premenných, ktoré spôsobujú náročnejšie vyjadrenie informácií pomocou vizualizácie [2].

1.2 Druhy vizualizácií

Informácie je možné zobrazovať v rôznych geometrických tvaroch aj negeometrických atribútoch vzhľadom, ako je napríklad farba alebo typ čiary. Vhodnosť výberu vizualizácie závisí od množstva a typu dát, ktoré chceme vizualizovať. Na dva odlišné typy informácií použijeme rôzne druhy vizualizácií, aby sme zabezpečili vysokú výpovednú hodnotu. Ak zobrazujeme plán aktivít na nasledujúci mesiac, tak by bolo vhodné využiť Ganttov diagram. Na vizualizáciu počtu absolventov FMFI v jednotlivých rokoch nám postačí napríklad stĺpcový graf.

V literatúre [1] z roku 1985 uvádza Vaganyan sedem typov grafov, ktoré sa využívajú v manažmente. Ku každému druhu v danom rozdelení je ilustrovaných niekoľko praktických príkladov, ku ktorým vieme ponúknuť novodobé ekvivalenty. Príloha A obsahuje pôvodné vizualizácie informácií z knihy [1], a zároveň návrh vizualizácií podobnej informácie v súčasnosti.

1.2.1 Ilustračný graf charakteristík objektu

Ilustračné grafy charakteristík objektu opisujú jeden objekt a súvislosti medzi jeho jednotlivými komponentami. Môžeme sem zaradiť rôzne klasifikačné schémy, grafy štruktúr, technologické plány, vývojové diagramy, schémy informačného toku apod. Hlavným cieľom grafov tohto druhu je prehľadné zobrazenie štruktúry objektu, vzájomnej polohy, postupnosti objektu a súvisiacich prvkov, či naznačenie smeru posúvania informácií [1].

V prípade klasifikačnej schémy je nevyhnutné si na začiatku určiť charakteristické znaky, podľa ktorých klasifikujeme objekt. Na priloženej vizualizácii Obr. A1 a k nej vytvorenej novodobej vizualizácii Obr. A2 je zobrazená štruktúra vzdelávacieho systému, kde berieme za klasifikačné kritérium vek žiaka a ročník, v ktorom by sa mal v príslušnom veku nachádzať. Pri schematickom znázornení sídliska, Obr. A3 a Obr. A4, využívame rôzne geometrické prvky na zakreslenie reálnych objektov. Príklad schémy informačného toku je ilustrovaný na Obr. A5 a na aktuálnejšom Obr. A6, kde je naznačený posun informácií v rámci jednotlivých oddelení podniku. Na vyobrazenie stavov spracovania dokumentu a jeho pohybu medzi poverenými oddeleniami Obr. A7 sa v súčasnosti môže využiť tabuľka MS Excel a v rámci nej napríklad funkcionálna podmieneného formátovania, Obr. A8.

Grafy, charakterizujúce jednotlivé objekty a jeho prvky, boli do novodobých softvérov implementované skôr ako kompaktné grafické schémy. V MS Office existuje funkcionálna SmartArt Graphic, ktorá dovoľuje používateľovi vytvoriť prehľadnú vizualizáciu rýchlo a jednoducho. Používateľ sa iba potrebuje rozhodnúť, akú informáciu chce zobrazit' – postupnosť procesov, hierarchickú štruktúru, zoznam krokov, SWOT analýzu v matici alebo niečo iné.

1.2.2 Ilustračný graf funkčných závislostí

Ilustračné grafy funkčných závislostí medzi jednotlivými premennými sa využívajú na analýzu určitého javu v plánovacom období – počas štvrťroka, polroka alebo roka. Tieto vizualizácie uľahčujú manažérom robiť rozhodnutia, vytvárať normy, evidovať aktuálny stav [1].

Na priloženej vizualizácii Obr. A9 a jej novodobej verzii Obr. A10 je ilustrovaná funkčná závislosť dvoch premenných. Príklad funkčnej závislosti troch premenných je

zobrazený na Obr. A11, pričom v podobných vizualizáciách vytváraných v súčasnosti môžeme, podľa potreby, použiť aj farebnú škálu ako ďalší rozmer, Obr. A12. Pri zobrazovaní funkčnej závislosti dvoch premenných je tiež možné využívať farebné odlíšenie zobrazovaných prvkov. Na vizualizácii Obr. A13 je na jednotlivých osiach zobrazený čas a stupeň riadenia. Zvýraznená lomená čiara predstavuje aktuálny stav v podniku. Ekvivalent tejto vizualizácii sme vytvorili v MS Excel, Obr. A14, kde sú farebne vyčlenené jednotlivé sektory v riadení, avšak grafu chýba krivka reálneho stavu. Niekedy má manažér podniku k dispozícii natoľko špecifické dáta, že ich vizualizáciu je potrebné robiť v špeciálnom softvéri, ktorý je prispôsobený na prácu s daným typom údajov.

1.2.3 Porovnávacie diagramy

Porovnávacie diagramy slúžia na porovnanie štatistických súborov vzhľadom na nejaký ľubovoľne meniaci sa znak. Porovnáваме veľkosti jednotlivých ukazovateľov, napríklad obsah alebo objem geometrických útvarov. Do tejto kategórie patria stĺpcové, čiarové alebo pruhové grafy [1].

Stĺpcové grafy môžeme zobrazovať dvojrozmerné, ako na priložených vizualizáciách Obr. A15 a Obr. A16, ale aj v priestore, Obr. A17 a Obr. A18. Stĺpcové grafy zvyčajne používame, ak chceme porovnať jednotlivé hodnoty v určitom segmente alebo za zvolené obdobie. Napríklad v dvojrozmernom grafe Obr. A16, ktorý sme ponúkli ako ekvivalent ku grafu Vaganyana [1], sledujeme návštevnosť internetového obchodu podľa typu používateľa a zariadenia, z ktorého webstránku otvoril. Ďalej predkladáme rôzne alternatívy pruhového grafu. Klasické pruhové grafy sú ilustrované na Obr. A19 a Obr. A20. Príklady grafov s nadobudnutými zápornými hodnotami sú predstavené na Obr. A21 a Obr. A22. V prípade, že potrebujeme zobraziť rozdelenie váh jednotlivých segmentov, môžeme použiť skladaný pruhový graf zobrazený na škále 100%, Obr. A23 a Obr. A24. Posledným príkladom porovnávacích diagramov je koláčový diagram, Obr. A25 a Obr. A26. Ku každému sektoru priradíme jeho reálnu hodnotu, na základe ktorej sa vypočíta váha prislúchajúceho úseku v diagrame. Počet sektorov by nemal presahovať číslo päť, aby sa zachovali jednoznačné rozdiely medzi segmentami, ktoré porovnáваме.

Na príslušných vizualizáciách, ktoré obsahuje Príloha A, si môžeme všimnúť zásadný rozdiel medzi porovnávacími diagramami z osemdesiatych rokov a novodobými. V súčasnosti prevláda minimalistický vzhl'ad, ktorý zvyšuje prehľadnosť grafov. Zároveň sa technická úroveň softvérov na tvorbu vizualizácií zvýšila do takej miery, že grafické prevedenie diagramov je atraktívnejšie a poskytuje používateľovi rozličné možnosti zobrazenia informácie.

1.2.4 Dynamické diagramy

Medzi základné typy dynamických diagramov patria lineárne súradnicové grafy, grafy na logaritmickú sieť a koláčové grafy v polárnych súradniciach. Najčastejšie sa využívajú lineárne súradnicové grafy, ktoré vystihujú dynamiku ekonomických ukazovateľov [1].

Jedným príkladom čiarového grafu je vizualizácia Obr. A27 v prílohe a jej súčasný ekvivalent Obr. A28. Charakteristickým znakom tohto typu grafu je, že jeho osi predstavujú súradnicovú sústavu a lomené čiary zobrazujú zmenu jedného parametra v priebehu rokov, napríklad zmenu počtu vedeckých pracovníkov na univerzite podľa získaného titulu alebo zmenu počtu absolventov fakulty podľa študijného odboru. Dynamiku zmeny jednotlivých ukazovateľov veľmi dobre vystihujú aj stĺpcové grafy. Prostredníctvom nich môžeme vyjadriť zmenu jedného ukazovateľa, ako je ilustrované na Obr. A29 alebo na Obr. A30, kde zobrazujeme zmenu počtu študentov univerzity v časovom rade. Stĺpcovým grafom môžeme vyjadriť aj zmenu viacerých ukazovateľov, napríklad na Obr. A31 sú zobrazené štyri rozmery – rok merania, počet podnikov, ich obrat a priemerný počet pracovníkov na podnik. Ako ekvivalent tejto vizualizácii navrhujeme kombinovaný graf Obr. A32, ktorý s využitím duálnych osí dokáže prehľadne vizualizovať všetky štyri rozmery. Kruhový graf Obr. A33, predstavujúci zmenu jedného ukazovateľa počas daného časového obdobia vieme v súčasnosti nahradiť radarovým grafom Obr. A34 s podobnými vlastnosťami. Vizualizácia Obr. A35 a k nej ekvivalentná Obr. A36 slúži na zaznamenanie krátkodobých zmien ekonomických ukazovateľov, prípadne na zobrazenie ich percentuálnej miery kolísania.

Diagramy vyjadrujúce dynamiku zmeny ukazovateľov určitého objektu sa v priebehu rokov zásadne nezmenili. Princípy ich tvorby, použitia a následného čítania ostali podobné. Avšak, výraznú zmenu zaznamenalo najmä grafické prevedenie diagramov a možnosť používania farebnej škály na odlíšenie jednotlivých parametrov.

1.2.5 Štatistické mapy

Štatistické mapy zobrazujú štatistické údaje v geografickom prostredí. Delia sa na tri základné typy – mapy zobrazujúce priemerné pôsobenie javu v rámci územia, mapy zobrazujúce celkovú veľkosť javu v rámci územia, historicko-geografické mapy zobrazujúce dlhodobý presun javu v rámci územia [1].

Pomocou štatistickej mapy môžeme zaznamenávať početnosť výskytu určitého úkazu v danej oblasti, napríklad Obr. A37 predstavuje poľnohospodárske plochy, ktoré boli zasiate a k nemu aktuálnejší Obr. A38 ukazuje množstvo a rozmiestnenie krajov v rámci jednotlivých štátov. Ďalší druh mapy, Obr. A39, plní analyticko-štatistickú úlohu. Na mape je vizualizovaná informácia o spôsobe vykonávania príkazov v jednotlivých mestských častiach. Ku každej geografickej oblasti je pridelený stĺpcový graf, v ktorom sa vedie štatistika o celkovom počte vydaných pokynov, o počte splnených úloh načas alebo s omeškaním, či nesplnených vôbec. K tomuto typu mapy je možné vytvoriť v súčasnosti oveľa prepracovanejší ekvivalent Obr. A40. Vďaka interaktivite niektorých vizualizačných nástrojov je možné zaznamenať analyticko-štatistickú informáciu v informačnom štítku, a tak ponechať mape jednoduchosť a prehľadnosť. Existujú softvéry, ktoré umožňujú vytvoriť mapy spájaním bodov s určitou polohou, Obr. A41. Vďaka podobným schémam získavame informácie o najkratšej ceste v grafe, ktoré pomáhajú hlavne optimalizovaniu cestných komunikácií. V súčasnosti už sú vytvorené programy, ktoré umožňujú používateľovi vyhľadávať vzdialenostne alebo časovo najkratšie cesty medzi zvolenými bodmi, vrátane ďalších doplnujúcich funkcionalít, Obr. A42.

V štatistických mapách pozorujeme od ich vzniku významné zmeny. Nielenže sa vylepšuje ich vizuálna príťažlivosť, ale hľadajú sa aj spôsoby efektívneho podania informácií. Počas pridávania ďalších dát k zemepisnému údaju je potrebné myslieť na spôsob ich vzájomného zobrazovania, aby ostala mapa čitateľná a poskytovala vysokú informačnú hodnotu. Mnoho problémov sa aktuálne rieši práve vďaka možnosti interakcie používateľa s počítačovým programom.

1.2.6 Plánový graf

Plánové grafy majú široké využitie v operatívnom plánovaní, v plánovaní súčasných aj budúcich aktivít. Tieto grafy sú vhodné na kontrolu a hodnotenie napĺňania plánov podniku, a taktiež pomáhajú pri robení manažérskych rozhodnutí. Medzi plánové grafy zaraďujeme evidenčný a kontrolný diagram, Ganttov diagram, sieťový graf a iné [1].

Príloha A obsahuje jediný príklad plánového grafu, tzv. Ganttov diagram. Tento diagram je spojením pruhového grafu a kalendáru, ktorý slúži na zaradenie nevyhnutných aktivít do časového plánu, Obr. A43. V Ganttovom diagrame vidíme začiatok naplánovanej činnosti, jej trvanie aj predpokladaný koniec. Zásadným vylepšením oproti starším vizualizačným technikám je možnosť interaktivity v novodobých vizualizačných nástrojoch. Výborným príkladom je vizualizácia Obr. A44, ktorá okrem prehľadného časového plánu dovoľuje používateľovi sledovať podrobnosti ku každej činnosti. V informačnom štítku si vieme pozrieť odhadovaný počet dní na vykonanie úlohy, dátum jej začiatku alebo meno osoby zodpovednej za splnenie úlohy. Navyše, pridaním farebného filtra do grafu môžeme jednoducho rozlíšiť osoby, poverené jednotlivými úlohami.

1.2.7 Sieťové grafy

Sieťové grafy tvoria samostatný druh vizualizácií z dôvodu ich vysokej početnosti a rôznorodosti. V praxi sa používajú najmä kombinované grafy a slúžia na optimalizáciu manažérskych rozhodnutí [1].

Vaganyan uvádza príklady sieťových grafov v rozličných podobách – od tradičného zobrazenia cez zložitejší vzhľad, s vrcholmi umiestnenými buď diagonálne, alebo horizontálne. Jednotlivé formy grafov z knihy [1] sme spojili do jednej ilustrácie, Obr. A45. Využitie tohto typu sieťového grafu je rozšírené najmä v oblasti informatiky na schematické zakreslenie postupnosti vykonávaných aktivít, procesov. Dnes existuje relatívne veľa rôznych softvérov, v ktorých sa dajú kresliť sieťové diagramy. My sme na vytvorenie aktuálnejšej verzie použili funkciu „forceNetwork“ z knižnice „networkD3“ softvéru R [6]. Výsledný graf Obr. A46 sa vykresľuje v okne internetového prehliadača a má interaktívne vlastnosti po kliknutí myškou na daný vrchol.

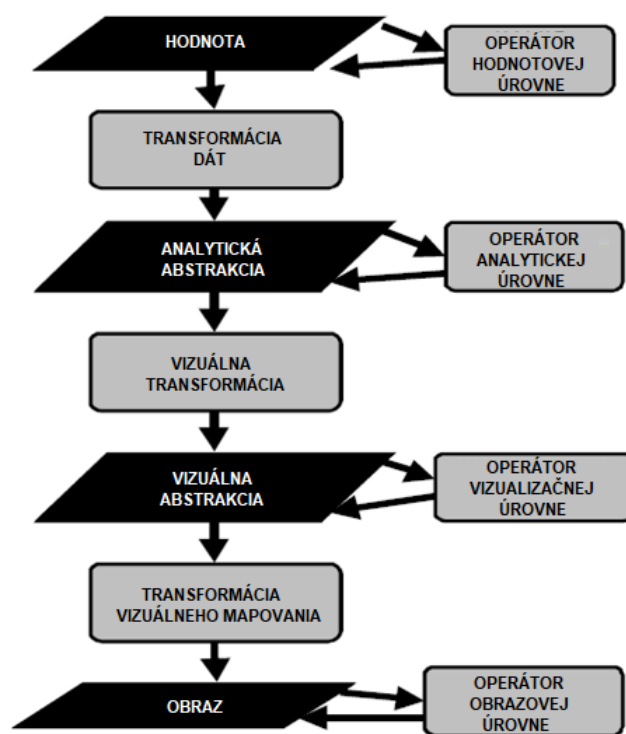
Na optimalizáciu manažérskych rozhodnutí je výhodné používať kombinované grafy, napríklad lineárne a sieťové. Podanie dát cez viacero druhov vizualizácií môže pomôcť manažérovi pri vytvorení si celkového obrazu o probléme a nájdení optimálneho riešenia. Ak by chcel manažér v praxi zlepšiť rozdelenie pracovnej sily, tak by bolo prospešné sa súčasne pozrieť napríklad na kalendárny graf trvania jednotlivých činností v podobe Ganttovho diagramu Obr. A44, graf dostupných ľudských zdrojov v danom období Obr. A10 a graf funkčnej závislosti ceny pracovníka od jeho času stráveného nad pridelenou činnosťou.

1.3 Referenčný model stavu informácií

Existuje viacero prístupov, akým sa dajú klasifikovať techniky vizualizácie informácií. Mnohé taxonómie majú v centre pozornosti práve dáta. Jedna z taxonómií, OLIVE, bola zostavená študentmi a delí techniky vizualizácie informácií podľa typu dát. Vizualizované dáta môžu byť ôsmich typov: časové, jednorozmerné, dvojrozmerné, trojrozmerné, viacrozmerné, strom, sieť, pracovný priestor [11].

Techniky vizualizácie môžu byť založené nielen na type informácií, ale aj na tzv. procesných operátoroch. Pomocou referenčného modelu stavu informácií vieme opísať rôzne techniky vizualizácie dát. Model je ilustrovaný na Obr. 3.

Každú techniku vizualizácie je možné rozdeliť na štyri dátové úrovne, na tri typy transformácií dát a na štyri operátory v rámci každej dátovej úrovne. Na hodnotovej úrovni sa nachádzajú surové dáta, t. j. namerané dáta bez akýchkoľvek úprav. Transformáciou dát sa z hodnoty vytvára určitá forma analytickej abstrakcie, zvyčajne vyťahovaním informácií zo surových dát. Na úrovni tejto analytickej abstrakcie získavame metadáta, t. j. doplňujúce dáta, ktoré opisujú surové dáta. Vizúálna transformácia ďalej redukuje metadáta na tie, ktoré je možné zobrazit', a teda vo vizuálnej abstrakcii sa nachádzajú také dáta, ktoré vieme vizualizovať pomocou skúmanej vizualizačnej techniky. Vizúálnym mapovaním sa premieňajú dáta, ktoré sa dajú zobrazit', priamo do grafickej interpretácie. Posledná obrazová úroveň ponúka konečný výsledok vizuálneho mapovania, kde pozorovateľ dokáže interpretovať informácie na základe obrázka [11].



Obr. 3 - Referenčný model stavu informácií [11]

Takto definovaná taxonómia vizualizačných techník slúži na pochopenie skúmanej techniky vďaka rozdeleniu kľúčových krokov v procese vizualizácie. Referenčný model stavu informácií je vhodný na kategorizovanie a klasifikovanie. Pomocou neho dokážeme odhaľovať vzťahy jednotlivých modulov vizualizácie, a zároveň skúmať spoločné a rozdielne znaky rôznych vizualizačných techník [11].

1.4 Paralelné súradnice

S pribúdajúcimi zdrojmi informácií pribúdajú aj nové vizualizačné techniky. V závislosti od počtu parametrov, na základe ktorých môžeme dáta merať, alebo od potrebných informácií pre konkrétny odbor, sa líši aj výber vhodnej techniky na vizualizovanie. Musíme mať jasno v tom, na aké otázky chceme vo vizualizácii nájsť odpoveď. Následne môžeme aplikovať jednu alebo viac vizualizačných metód.

V tejto súvislosti bola na akademickej pôde vo Švajčiarsku vytvorená vizualizačná metafora v podobe periodickej tabuľky. Ruský chemik Mendelejev predstavil v roku 1869

periodickú tabuľku chemických prvkov, v ktorej logicky zoradil prvky podľa ich vlastností. V jej premyslenom usporiadaní našli inšpiráciu autori odborného článku [12] a vytvorili tzv. periodickú tabuľku vizualizačných metód, Obr. 4. Tabuľka prezentuje dva rozmery – jednotlivé prvky a vizualizačné skupiny – a poskytuje prehľad užitočných vizualizačných metód, ktoré je možné uplatniť nielen v manažmente [12].

A PERIODIC TABLE OF VISUALIZATION METHODS

Data Visualization Visual representations of quantitative data in schematic form (either with or without axes)		Strategy Visualization The systematic use of complementary visual representations in the analysis, development, formulation, communication, and implementation of strategies in organizations.												G graphic facilitation																					
Information Visualization The use of interactive visual representations of data to amplify cognition. This means that the data is transformed into an image, it is mapped to screen space. The image can be changed by users as they proceed working with it.		Metaphor Visualization Visual Metaphors position information graphically to organize and structure information. They also convey an insight about the represented information through the key characteristics of the metaphor that is employed.		Concept Visualization Methods to elaborate (mostly) qualitative concepts, ideas, plans, and analyses.		Compound Visualization The complementary use of different graphic representation formats in one single schema or frame.		Me meeting trace		Mm metro map		Tm temple		St story template		Tr tree		Ct cartoon																	
Pi pie chart		L line chart						Go communication diagram		Fp flight plan		Cs concept skeleton		Br bridge		Fu funnel		Ri rich picture																	
B bar chart		Ac area chart		R radar chart cobweb		Pa parallel coordinates		Hy hyperbolic tree		Cy cycle diagram		T timeline		Ve vean diagram		Mi mindmap		Sq square oppositions		Cc concentric circles		Ar argument slide		Sw swim lane diagram		Gc gantt chart		Pm perspectives diagram		D dilemma diagram		Pr parameter ruler		Kn knowledge map	
Hi histogram		Sc scatterplot		Sa sankey diagram		In information lense		E entity relationship diagram		Pt petri net		Fl flow chart		Cl clustering		Lc layer chart		Py pyramid technique		Ce cause-effect chains		Tl toulmin map		Dt decision tree		Cp cpm critical path method		Cf concept fan		Co concept map		Ic iceberg		Lm learning map	
Tk tokay box plot		Sp spectrogram		Da data map		Tp treemap		Cn cone tree		Sy system dyn./simulation		Df data flow diagram		Se semantic network		So soft system modeling		Sn synergy map		Fo force field diagram		Ib ibo argumentation map		Pr process event chains		Pe petri chart		Ev evocative knowledge map		V vee diagram		Hh heaven's hell chart		I informal	

Note: Depending on your location and connection speed it can take some time to load a pop-up picture. version 1.5
© Ralph Lengler & Martin J. Eppler, www.visual-literacy.org

Cy Process Visualization		Hy Structure Visualization		Su supply demand curve		Pc performance charting		St strategy map		Oc organisation chart		Ho house of quality		Fd feedback diagram		Ft failure tree		Mq magic quadrant		Ld life-cycle diagram		Po porter's five forces		S s-cycle		Sm stakeholder map		Is ishikawa diagram		Tc technology roadmap	
Ed edgeworth box		Pf portfolio diagram		Sg strategic game board		Mz mintzberg's organigraph		Z zwick's morphological box		Ad affinity diagram		De decision discovery diagram		Bm bcg matrix		Stc strategy canvas		Vc value chain		Hy hype-cycle		Sr stakeholder rating map		Ta taps		Sd spray diagram					

Obr. 4 - Periodická tabuľka vizualizačných metód [12]

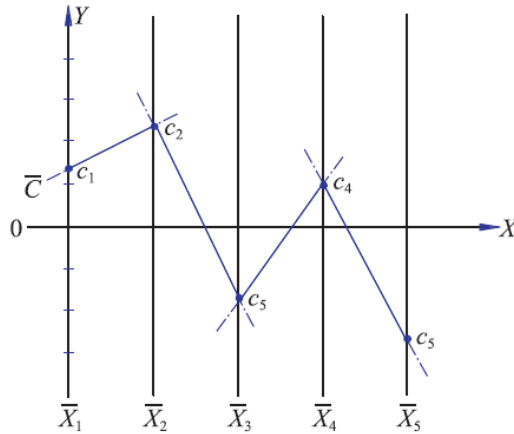
Na internetovej stránke [13] je zakomponovaná do tejto vizualizačnej metafory aj interakcia, ktorá umožňuje používateľovi prezerat' si konkrétne príklady prezentovaných techník. Vďaka rozdeleniu metód na skupiny vieme spozorovať, že čím viac sledovaných premenných sa v dátach nachádza, tým viac sa zvyšuje zložitosť potrebnej vizualizačnej metódy. Dvojmerné dáta je možné zobraziť v rovine, napríklad ako histogram alebo čiarový graf. Trojmerné dáta si dokážeme predstaviť v priestore. Pri štvormerných dátach môžeme pridať do priestoru napríklad farebnú škálu ako štvrtý rozmer. Avšak, vizualizácia päť- a viacmerných informácií už môže predstavovať náročnejší problém.

Vizualizovať viacrozmerné dáta nie je jednoduché, ak chceme zachovať prehľadnosť a poskytnúť pozorovateľovi potrebný nadhľad nad problémom. Podľa počtu premenných sa vyberá vhodná vizualizačná metóda. Ak sa chceme dištancovať od výberu metódy na základe počtu zobrazovaných rozmerov, tak jedným z riešení by mohlo byť využitie paralelných súradníc.

Metóda paralelných súradníc je založená na princípe dvojrozsmernej karteziánskej sústavy súradníc a rovnobežnosti. Touto témou sa začal zaoberať Alfred Inselberg už ako študent v roku 1959. Jeho cieľom bolo dosiahnuť systematickú vizualizáciu viacrozmerných problémov bez straty informácie [2]. Vo svojej publikácii [2] uvádza päť základných vlastností, ktoré by mala takáto vizualizačná metóda spĺňať:

1. Metóda má nízku výpočtovú zložitosť, napr. $O(n)$ pre n premenných, a je lineárna v počte rozmerov, resp. premenných.
2. Metódu je možné aplikovať pre ľubovoľnú dimenziu a s každou premennou sa zaobchádza rovnako.
3. Zobrazené objekty je možné rozpoznať aj po uskutočnení transformácie ako napríklad otočenie, posunutie, úprava mierky.
4. Vlastnosti n -rozmerných objektov sú zrejmé zo zobrazenia.
5. Metóda je matematicky presná.

Inselberg vo svojom diele [2] vysvetľuje princíp paralelných súradníc. Nech je E^2 dvojrozmerný euklidovský metrický priestor s karteziánskymi súradnicami x a y . Kolmo na x -ovú os rovnomerne umiestnime n kópií reálnej priamky \mathbb{R} , označíme ich $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n$. Vytvorili sme osi paralelnej súradnicovej sústavy pre n -rozmerný euklidovský priestor E^n . Bod $C \in E^n$ so súradnicami (c_1, c_2, \dots, c_n) je v paralelných súradniciach zobrazený ako lomená čiara \bar{C} , pričom n vrcholov tejto lomenej čiary leží v bode $(i-1, c_i)$ na osi \bar{x}_i pre $i = 1, \dots, n$. Lomená čiara \bar{C} je kompletná, teda obsahuje úsečky medzi jednotlivými osami. Takýmto spôsobom sme zabezpečili korešpondenciu medzi bodmi v n -rozmernom euklidovskom priestore E^n a rovinnými lomenými čiarami s vrcholmi ležiacimi na rovnobežných osiach [2]. Príklad paralelnej súradnicovej sústavy pre $n = 5$ je uvedený na Obr. 5.



Obr. 5 - Lomená čiara \bar{C} predstavuje bod $C = (c_1, c_2, c_3, c_4, c_5)$ v paralelných súradniciach [2]

1.5 Otázky vizuálneho myslenia

Pred vytváraním vizualizácií je nevyhnutné si uvedomiť, k akým problémom chceme získať riešenie. Najprv by sme si mali položiť rozumnú otázku a následne na ňu hľadať odpoveď. Dan Roam vo svojej knihe [14] opisuje šesť základných otázok v jazykovej komunikácii a tvrdí, že každý problém, vzniknutý z niektorej základnej otázky, je možné vyriešiť nakreslením obrázka. Nasledujúcich šesť otázok (z angl. 6W) nám pomáha vidieť problém a určiť jeho typ a obrázok k riešeniu – kto alebo čo (z angl. who, what), koľko (z angl. how much), kde (z angl. where), kedy (z angl. when), ako (z angl. how) a prečo (z angl. why).

Je dôležité pochopiť, ako problém vidíme, aby sme sa vedeli vhodne opýtať pred hľadaním riešenia. Ku každej otázke 6W prislúcha práve jeden spôsob ukazovania, ku ktorému existuje vizuálna schéma. Ak vidíme problém, na ktorý sa pýtame kto alebo čo, tak riešenie ukazujeme vo forme portrétu (z angl. portrait). Na otázku koľko je možné odpovedať grafom, diagramom (z angl. chart). Pri otázke kde ukazujeme mapu (z angl. map), pri otázke kedy ukazujeme časovú os (z angl. timeline), pri otázke ako ukazujeme sieťový graf (z angl. flowchart) a na otázku prečo reagujeme viacrozmerným grafom (z angl. multiple variable plot) [14]. Roam vytvoril rámec vizuálneho myslenia, v ktorom objasňuje vznik súradnicových systémov zo šiestich otázok 6W.

		② → which version	S. <u>simple</u> elaborate	Q. <u>quality</u> quantity	V. <u>vision</u> execution	I. <u>individual</u> comparison	Δ. <u>change</u> as-is
① which framework ↓							
1 who/what? (portrait)		↔					
2 how much? (chart)		↔					
3 where? (map)		↔					
4 when? (timeline)		↔					
5 how? (flowchart)		↔					
6 why? (plot)		↔					

Obr. 6 - Kódex vizuálneho myslenia [14]

Ďalšou pomôckou, ktorú Roam navrhol na tvorbu vizualizácie s vysokou výpovednou hodnotou pre daného pozorovateľa, je metóda SQVID. SQVID je séria piatich otázok, ktorými by sme mali previesť prvotný nápad, aby sme zabezpečili jeho vizuálnu zrozumiteľnosť

a upresnili myšlienku. V závislosti od toho, ako chceme ukazovať riešenie problému, tak sa v každej otázke rozhodujeme medzi dvoma možnosťami – má byť obrázok jednoduchý (z angl. simple) alebo rozpracovaný (z angl. elaborate), ukazujeme kvalitu (z angl. quality) alebo kvantitu (z angl. quantity), víziu (z angl. vision) alebo realizáciu (z angl. execution), kladieme dôraz na jednotlivé atribúty (z angl. individual) alebo porovnávame (z angl. comparison), chceme znázorniť zmenu (z angl. change) alebo status quo (z angl. as-is) [14].

Roam vytvoril dva spôsoby, akým je možné ukazovať riešenie problému. Jeho rámec vizuálneho myslenia podporuje analytické myslenie a model SQVID zase intuitívne myslenie. Vďaka tomu, že oba spôsoby sa navzájom veľmi líšia, tak ich rozdiely dodávajú modelom komplementárnosť. Ak sa využijú súčasne, tak riešenie problému zbadáme takmer okamžite. Roam umiestnil tieto dva modely do súradnicového systému, a vytvoril efektívny nástroj na riešenie problémov pomocou obrázkov, tzv. kódex vizuálneho myslenia, Obr. 6 [14].

1.5.1 Štatistické pojmy

Vhodnosť a aktuálnosť kladených otázok pri analyzovaní vizualizácií závisí aj od povahy informácií, ktoré chceme zistiť. V diplomovej práci budeme využívať poznatky z počítačovej, finančnej, matematickej či štatistickej oblasti na vytváranie niektorých grafov. Uvedieme si niekoľko základných pojmov zo štatistiky, ktoré sa neskôr vyskytnú v práci. Definície sú prevzaté hlavne z literatúry [15].

„O vektore $X = (X_1, X_2, \dots, X_n)'$ hovoríme, že má prvé momenty, ak existujú stredné hodnoty jeho zložiek $E(X_1), E(X_2), \dots, E(X_n)$ a výraz $E(X) = (E(X_1), E(X_2), \dots, E(X_n))'$ nazývame jeho strednou hodnotou.“ [15]

Stredná hodnota je špeciálnym prípadom skupiny parametrov – momentov. Zaradíme ich medzi miery polohy, rovnako ako aj kvantily. Medzi kvantily patria percentily, decily a najčastejšie využívané kvartily. Medián je tiež kvantil, a teda je mierou polohy.

Na rozdiel od aritmetického priemeru, tak medián nazývame robustnou štatistikou. Robustné metódy, resp. štatistiky, je vhodné použiť, keď očakávame výskyt odľahlých hodnôt, t. j. extrémnych hodnôt veľmi vzdialených od strednej hodnoty. Výhodou robustných štatistík

je, že sú odolné voči odľahlým hodnotám, a vedia s nimi pracovať bez toho, aby bol výsledok výpočtu zásadne ovplyvnený [16].

„Nech $\mathbf{X} = (X_1, X_2, \dots, X_n)'$ má konečné druhé momenty, $E(X_i^2) < \infty, i = 1, 2, \dots, n$. Potom kovarianciou premenných X_i, X_j pre $1 \leq i, j \leq n$ budeme nazývať výraz $cov(X_i, X_j) = E\{(X_i - E(X_i))(X_j - E(X_j))\}$.“ [15]

Z praktických dôvodov budeme využívať zápis kovariancie v tvare $cov(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j)$. Medzi vlastnosti kovariancie premenných X_i, X_j patrí prípad, keď $cov(X_i, X_j) = 0$. Vtedy hovoríme, že X_i a X_j sú nekorelované náhodné premenné [15].

„Nech $\mathbf{X} = (X_1, X_2, \dots, X_n)'$ je náhodný vektor. Nech pre disperzie náhodných premenných $X_k, k = 1, 2, \dots, n$ platí $D(X_k) < \infty$. Kovariančnou maticou náhodného vektora \mathbf{X} nazývame symetrickú $n \times n$ rozmernú maticu, ktorá má na priesečníku i -tého riadku a j -tého stĺpca číslo $cov(X_i, X_j), i, j = 1, 2, \dots, n$; t.j. $\mathcal{D}(\mathbf{X}) = (\sigma_{ij})$, kde $(\sigma_{ij}) = cov(X_i, X_j)$.“ [15]

„Nech X, Y sú náhodné premenné, pre ktoré platí $E(X^2) < \infty, E(Y^2) < \infty, D(X) > 0, D(Y) > 0$. Potom číslo $\rho_{X,Y} = \frac{cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$ nazývame korelačným koeficientom (koeficientom korelácie) náhodných premenných X a Y .“ [15]

„Maticu $\mathcal{R}(\mathbf{X})$, ktorá má na priesečníku i -tého riadku a j -tého stĺpca číslo ρ_{X_i, X_j} , nazývame korelačnou maticou náhodného vektora $\mathbf{X} = (X_1, X_2, \dots, X_n)'$.“ [15]

Koreláciu medzi premennými je možné počítat' viacerými metódami. Pri výbere vhodnej metódy je dôležité vedieť, či dáta pochádzajú z normálneho rozdelenia [15]. Overenie normality dát urobíme potvrdením alebo vyvrátením zostrojenej hypotézy. Štatistickou hypotézou nazývame „tvrdenie, ktoré sa týka rozdelenia pravdepodobnosti náhodných premenných alebo hodnôt parametrov týchto rozdelení“ [15]. Nulovou hypotézou H_0 nazývame tvrdenie, ktorého správnosť overujeme pomocou testovania štatistických hypotéz. Oproti nulovej hypotéze kladieme alternatívnu hypotézu H_1 .

Overenie hypotézy, či dáta pochádzajú z normálneho rozdelenia je možné urobiť cez Shapiro-Wilkov test [17]. Pre malé výbery premenných sú jeho výsledky spoľahlivé, preto tento test nachádza časté uplatnenie v počítačovej štatistike. Pri testovaní štatistických hypotéz

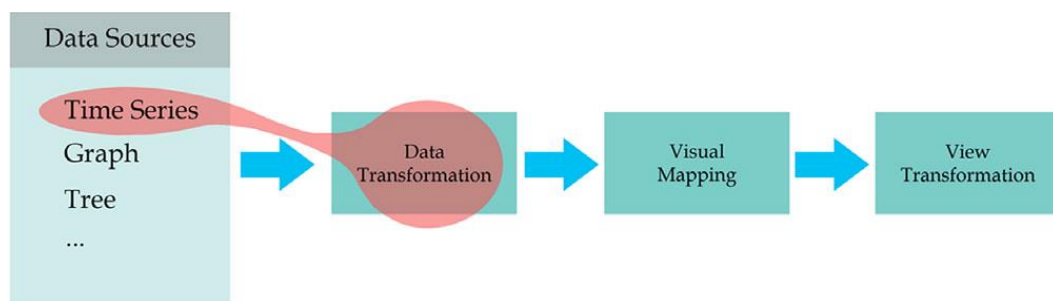
budeme zamietat' nulovú hypotézu, ak vypočítaná hodnota pravdepodobnosti je menšia ako určená kritická hodnota.

1.5.2 Časové rady

Častým prípadom vo vizualizácii informácií je zobrazovanie časových radov. Keďže nachádzajú široké uplatnenie v rôznych vedných disciplínach, ako napríklad vo finančnej oblasti, vo vede či v medicíne, tak je dôležité ich vedieť graficky reprezentovať. Vo všeobecnosti, časové rady sú jedným z typov dát a sú prezentované usporiadanou postupnosťou pozorovaní. Pozorovanie zvyčajne predstavuje konkrétnu hodnotu jedného atribútu, alebo kombináciu hodnôt viacerých atribútov [18].

Na základe hodnoty, ktorú časový rad nadobúda, rozlišujeme jednorozmerné a viacrozmerné časové rady. Príkladom viacrozmerých časových radov môže byť skupina viacerých parametrov o počasí, v ktorej meriame teplotu, rýchlosť vetra, smer vetra a zrážky, ale každému z týchto parametrov je priradená práve jedna premenná. Ak pre rovnaké parametre merania zaznamenávame viacero hodnôt, napríklad teplotu v rôznych mestách, tak hovoríme o viacrozmerých časových radoch so spoločnou premennou [18]. Jedná sa o dva rôzne prípady, ktoré si vyžadujú odlišný prístup počas analýzy a analytickej abstrakcie.

Shurkhovetskyy vo svojom výskume [18] podrobnejšie rozoberá spôsob transformácie časových radov na úroveň analytickej abstrakcie. Na ilustrovanie obsahu svojej práce uviedol Obr. 7 s vyznačenými témami. Tento obraz predstavuje proces vizualizácie časových radov, odvodený z referenčného modelu stavu informácií [11].



Obr. 7 - Hlavné úrovne v procese vizualizácie informácií [18]

2 Špecifikácia praktickej časti

Medzi veľmi účinné spôsoby získavania poznatkov a učenia sa novým veciam je skúšať teóriu v praxi. Z tohto dôvodu sme sa rozhodli spojiť teoretickú časť o vizualizácii informácií s praktickou časťou. V tejto kapitole predstavíme čitateľovi dáta z finančnej oblasti, ktoré využívame v diplomovej práci na vysvetlenie jednotlivých tém a na tvorbu vlastných vizualizácií. Taktiež venujeme pozornosť aj nástrojom, pomocou ktorých sme vlastné vizualizácie vytvorili.

Cieľom praktickej časti je snaha o vytvorenie paralelnej súradnicovej sústavy vo vizualizačnom softvéri, v ktorej bez straty informácií dokážeme vizualizovať viacrozmerné dáta. Zároveň sa snažíme prepojiť teóriu s praxou a ukázať čitateľovi spôsoby, ako získavať potrebné informácie z vizualizácií.

2.1 Dáta o Bitcoine

Pre praktickú časť diplomovej práce sme si potrebovali vybrať dáta, s ktorými budeme pracovať. V súčasnosti sa rozrástá trend ťažby a využívania kryptomien, vznikajú kurzy zamerané na výučbu o technológii blockchain a o princípe fungovania kryptomien, napríklad University of Nicosia ponúka vzdelanie v tejto oblasti [19]. Kryptomeny sa postupne stávajú alternatívou k peňažným platbám, a má zmysel si o nich zisťovať viac informácií a rozširovať si prehľad.

Prvou decentralizovanou kryptomenou je Bitcoin, ktorý aktuálne patrí aj medzi najpopulárnejšie kryptomeny v obeh. Jeho cena na konci roku 2017 výrazne vzrástla a neskôr začala opäť prudko klesať. Myslíme si, že by mohlo byť zaujímavé zisťovať dôvody a okolnosti takéhoto vývoja na základe vizualizácií, rovnako ako aj pokúsiť sa o predpoveď budúceho vývoja. Z týchto dôvodov sme sa rozhodli pre prácu s dátami o kryptomene Bitcoin.

Zdrojom dát je prevažne webová stránka CoinMarketCap [3], ktorá zhromažďuje informácie o kryptomenách, vytvára z nich vlastné vizualizácie či vedie štatistiky. Na vizualizovanie a analýzu sme si zvolili časový rad dostupných hodnôt od 16. januára 2017 do 15. januára 2019, pretože v tomto období bol vývoj Bitcoinu najzaujímavejší. Do stĺpca A dátového súboru uloženého v MS Excel sme doplnili identifikačné číslo riadku pre prípad, že

by sme ho chceli využívať pri identifikácii záznamu namiesto použitia dátumu. Okrem časovej zložky v stĺpci B obsahuje dátový súbor aj informáciu o trhovej cene v amerických dolároch. Uvádzame počiatočnú cenu v stĺpci C, najnižšiu v stĺpci D, najvyššiu v stĺpci E a uzatváraciu cenu v stĺpci F, za ktorú sa Bitcoin v daný trhový deň obchodoval. V stĺpci G sme si sami dopočítali percentuálny nárast, resp. pokles uzatváracej ceny daného trhového dňa oproti predchádzajúcemu. Ďalej máme zo stránky CoinMarketCap [3] v stĺpci H údaje o obchodovanom objeme v amerických dolároch, t. j. celkový objem transakcií – predajov a nákupov – vykonaných nad Bitcoinom v daný trhový deň. Jedine údaje o počte Bitcoinov v stĺpci I, ktoré sa nachádzali v obehú za daný trhový deň v priebehu zvolených dvoch rokov, sú z webovej stránky Blockchain [4]. Zvyšné dáta v dátovom súbore sme získali tiež z webu CoinMarketCap [3]. Nachádza sa tam údaj o maximálnom množstve Bitcoinov v stĺpci J, ktoré môže byť vyťažené počas celej jeho životnosti. Toto maximum je fixne stanovené na dvadsaťjeden miliónov Bitcoinov. Stanovením trhovej kapitalizácie v stĺpci K sa ponúka jedna z možností, ako určiť relatívnu hodnotu kryptomeny. Je to teda množstvo fiat meny prepočítanej na americký dolár, ktorá bola investovaná do Bitcoinu. Trhová kapitalizácia sa počíta ako súčin celkového počtu Bitcoinov v obehú a trhovej, uzatváracej, ceny Bitcoinu. Tento údaj je prezentovaný v amerických dolároch. Posledným údajom v našom dátovom súbore, v stĺpci L, je dominancia kryptomeny Bitcoin spomedzi ostatných kryptomien sledovaných na CoinMarketCap [3]. Dominancia sa určuje na základe trhovej kapitalizácie každej kryptomeny a vyjadruje sa v percentách. Ukážka z kompletného dátového súboru je zobrazená na Obr. 8.

	A	B	C	D	E	F	G	H	I	J	K	L
1	ID	Date	Open [USD]	High [USD]	Low [USD]	Close [USD]	Difference [%]	Volume [USD]	Circulating Supply [BTC]	Max Supply [BTC]	Market Cap [USD]	Dominance [%]
2	1	16.01.2017	821.78	834.53	820.27	831.53	1.17	82755200	16106863	21000000	13393372620	84.91
3	2	17.01.2017	830.95	910.56	830.8	907.94	8.42	155095008	16108838	21000000	14625779851	85.30
4	3	18.01.2017	909.37	917.5	858.3	886.62	-2.40	225676992	16110713	21000000	14284036613	85.43
5	4	19.01.2017	888.34	904.61	884.34	899.07	1.38	105625000	16112888	21000000	14486616700	85.31
6	5	20.01.2017	898.17	899.4	887.01	895.03	-0.45	86728400	16115250	21000000	14423545371	85.03
7	6	21.01.2017	895.55	927.37	895.53	921.79	2.90	111158000	16117513	21000000	14856911163	85.32
397	396	15.02.2018	9488.32	10234.8	9395.58	10166.4	6.61	9062540288	16867188	21000000	171477807437	34.96
398	397	16.02.2018	10135.7	10324.1	9824.82	10233.9	0.66	7296159744	16869150	21000000	172637061144	35.65
399	398	17.02.2018	10207.5	11139.5	10149.4	11112.7	7.91	8660880384	16871013	21000000	187482083882	36.12
400	399	18.02.2018	11123.4	11349.8	10326	10551.8	-5.32	8744009728	16873038	21000000	178040648022	36.73
727	726	11.01.2019	3674.02	3713.88	3653.07	3687.37	0.23	5538712865	17476813	21000000	64443301117	52.22
728	727	12.01.2019	3686.97	3698.98	3653.81	3661.3	-0.71	4778170883	17478538	21000000	63994140882	52.37
729	728	13.01.2019	3658.87	3674.76	3544.93	3552.95	-3.05	4681302466	17480250	21000000	62106461671	52.35
730	729	14.01.2019	3557.31	3727.84	3552.29	3706.05	4.13	5651384490	17482113	21000000	64789619995	52.72
731	730	15.01.2019	3704.22	3720.15	3619.95	3630.68	-2.08	5537192302	17483763	21000000	63477817959	52.32
732												

Obr. 8 - Ukážka z dátového súboru o Bitcoine

Zo zhromaždených dát o Bitcoine vieme zistiť minimálne päť rozmerov. Ak neberieme do úvahy konštanty, akou je maximálny počet Bitcoinov, ani dáta odvodené zo vzájomných

výpočtov, či hodnoty vyjadrujúce tú istú veličinu, akou je trhova cena, tak mozeme uvazovat o nasledujucich datach ako o jedinecnych: datum (ozn. B), uzatvaracia cena (ozn. F), objem transakci v dolaroach (ozn. H), pocet Bitcoinov v obehu (ozn. I) a dominancia spomedzi inych kryptomien (ozn. L). Tieto udaje mozeme neskor pouzit aj na vizualizaci pomocou metody paralelnych suradnic [2].

2.2 Vizualizacny nastroj Tableau

Kazdorocne sa vo svete zvyšuje tvorba dat a pribudaju nové zdroje informaci. Tato skutocnost moze znamenat pre ludstvo rychlejší rozvoj a napredovanie v technologiach. Avsak, na další vyvoj vo vede a biznise je potrebne rozumiet ziskanym datam.

So vznikom informacnych technologi vznikol priestor na vytvaranie roznych softverov, ktoré dokazu spracovavat informacie. Jednym z nich je aj nastroj Tableau, ktorý sluzi na vizualizaci dat [5]. Cielom Tableau je, aby ludia boli schopni vidiet a rozumiet datam, dokazali riesit problemy na zaklade dat, a zaroven si rozvijali svoje analyticke schopnosti. Prostredie softveru Tableau je jednoduche na pouzvanie, prehladne, a vysledky analyz su vkusne a prakticky vizualizované. Navyse, poskytuje rocnu študentsku licenciu na vyuzvanie softveru.

Softver Tableau sme si vybrali na tvorbu a analyzu vlastnych vizualizaci potrebnych k diplomovej praci. Zaroven chceme nastroj vyuzit na vytvorenie vizualizacie nami zvolenych dat v paralelnych suradniciach, ktoré sa nenachadzaju v predvolenom vybere softveru.

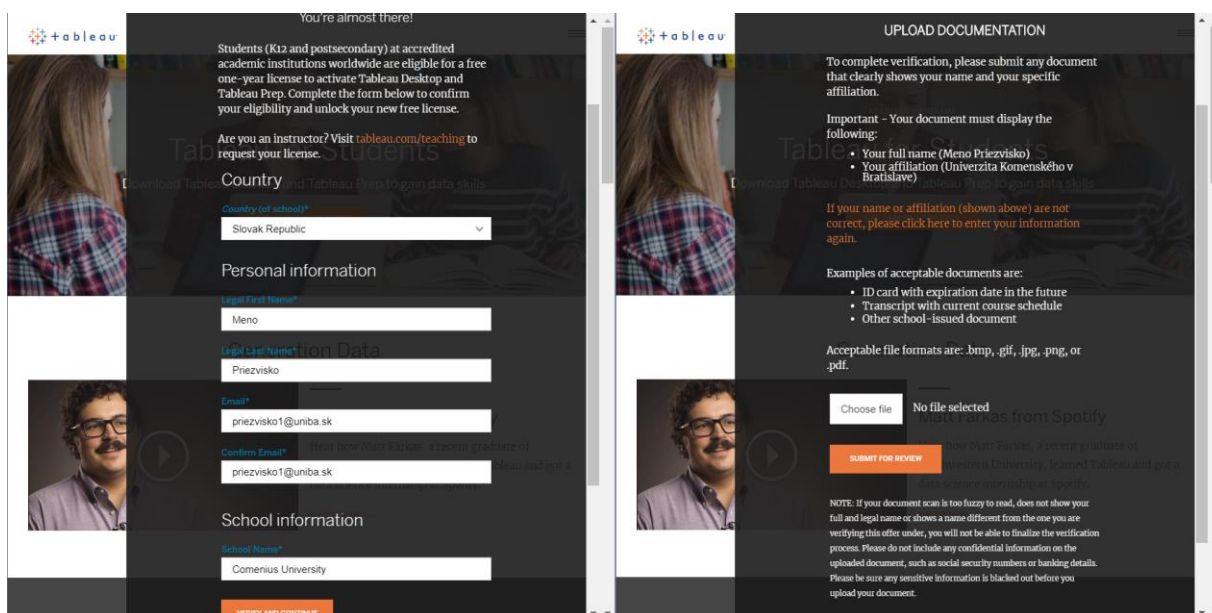
2.2.1 Ziskanie študentskej licencie

Na ziskanie rocnej študentskej licencie na vizualizacny nastroj Tableau je potrebne otvorit webovu stranku [5] so sekciou pre študentov a stlacit tlačidlo „GET TABLEAU FOR FREE“.

Na obrazovke sa otvori registracny formular, kde je potrebne vybrat zo zoznamu krajiny „Slovak Republic“, vyplnit meno, priezvisko a univerzitny email študenta, a napisat nazov univerzity „Comenius University“. Po stlaceni tlačidla „VERIFY AND CONTINUE“ sa overi spravnost a jedinecnost zadanych udajov. Na dalšej obrazovke registracneho formulara je

potrebné nahrať fotografie oboch strán aktuálnej študentskej karty ISIC, s čitateľným menom študenta, názvom univerzity a dátumom expirácie študentskej karty. Proces registrácie je ilustrovaný na Obr. 9.

Po stlačení tlačidla „SUBMIT FOR REVIEW“ je formulár odoslaný na overenie, ktoré môže trvať dlhší čas. V prípade úspešného overenia príde študentovi na zadaný univerzitný email potvrdenie o schválení študentskej licencie a aktivačný kód k softvéru, s platnosťou jeden rok.



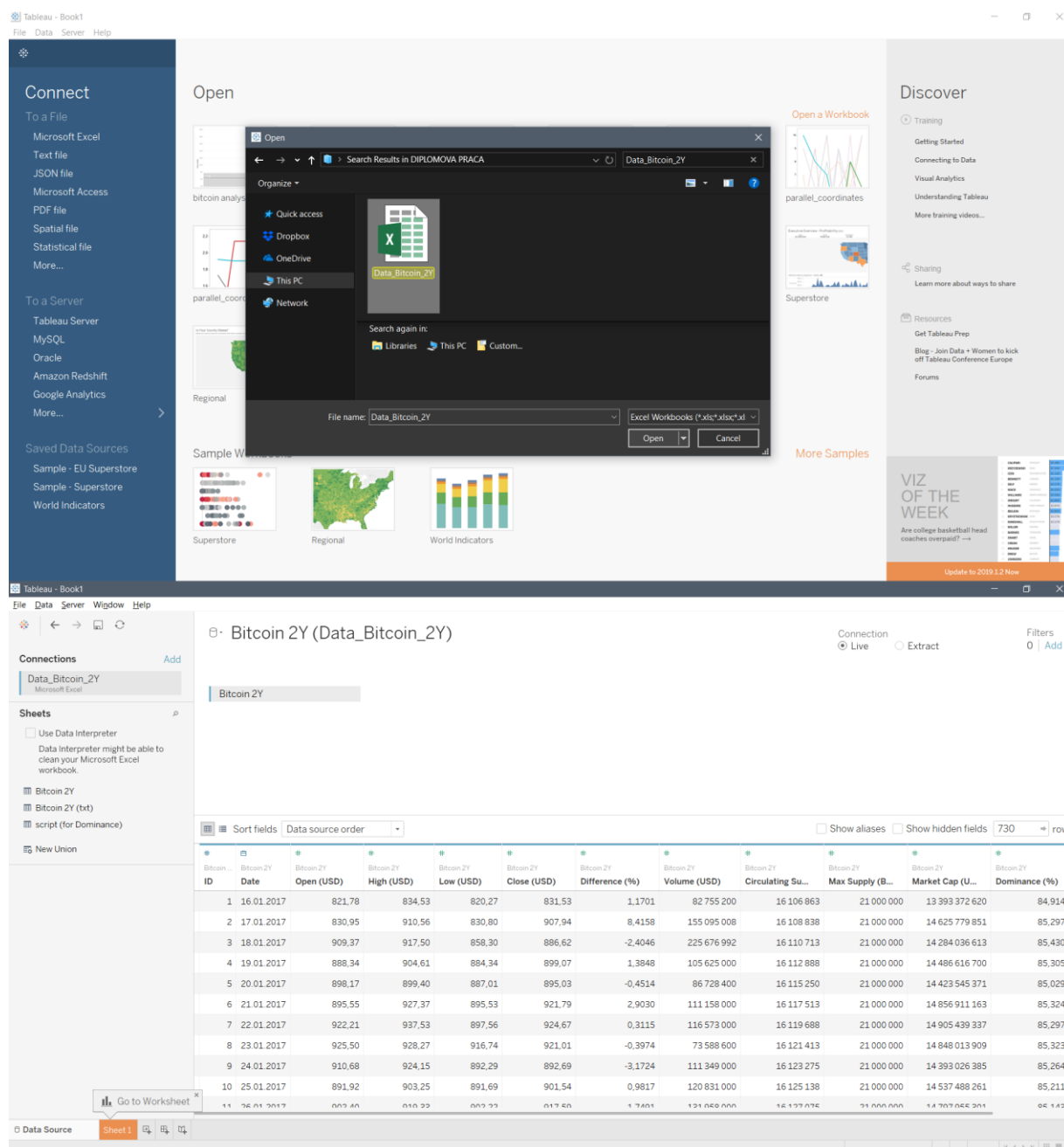
Obr. 9 - Registračný formulár na www.tableau.com/academic/students

2.2.2 Prvý kontakt s Tableau Desktop

Po získaní aktivačného kódu k softvéru Tableau sa študentovi ponúka možnosť na využívanie dvoch typov aplikácií. Produkt Tableau Prep slúži skôr na prípravu komplexnejších dát k analyzovaniu. Produkt Tableau Desktop je vhodný na prácu s upravenými dátovými súborami a jednoduché vytváranie vizualizácií.

My budeme v praktickej časti používať Tableau Desktop, keďže náš dátový súbor nepotrebuje ďalšie úpravy. Po otvorení počítačového programu si používateľ musí zvoliť zdroj dát, ktorý bude počas práce využívať. Ak sa dáta nachádzajú v súbore MS Excel, tak v ľavej lište „Connect“ v časti „To a File“ si zvolí „Microsoft Excel“ a vyberie potrebný dátový súbor.

Po pripojení tabuľky na pracovný priestor si používateľ zvolí hárky, ktoré chce využívať. S hárkami je možné robiť aj logické operácie ako napr. zjednotenie či prienik, ak si to spracovanie dát vyžaduje. Po nevyhnutnej úprave dát v dátovom zdroji môže používateľ prejsť do pracovného prostredia Tableau a vytvárať prvé analytické vizualizácie. Na Obr. 10 je ilustrovaný postup pridania dátového zdroja po otvorení softvéru Tableau.



Obr. 10 - Pridanie dátového zdroja v softvéri Tableau

2.3 Štatistický nástroj R

Často používaným softvérom medzi študentmi FMFI UK je R [6]. Praktické využitie nachádza najmä v spojení so štatistickými predmetmi. Softvér R je voľne dostupný štatisticko-analytický program, v ktorom môže používateľ robiť matematické výpočty, vykonávať rozličné štatistické operácie, testovať hypotézy či vytvárať jednoduché vizualizácie.

Výhodou softvéru R je, že má k dispozícii široký výber knižníc, vylepšujúcich balíčkov, s jednoduchým procesom inštalácie. Ak používateľ potrebuje v programe R rozšíriť poskytované funkcie, môže si priamo cez grafické používateľské rozhranie nainštalovať potrebný balíček. Napríklad na tvorbu vizualizácií v priestore je vhodné použiť funkcie z knižnice „plot3D“.

Rozhodli sme sa využiť známe prostredie nástroja R na získanie a overenie niektorých výpočtov pri dátových analýzach, na vytvorenie vizualizácií v priestore, keďže softvér Tableau ich neposkytuje v štandardnej ponuke, a na prípadné porovnanie vizualizácií s vizualizáciami vytvorenými v softvéri Tableau.

3 Vizualizácie v praxi

V praktickej časti diplomovej práce sa zoznámime s rôznymi druhmi vizualizácií, ktoré aplikujeme na vybrané dáta o Bitcoine. Teóriu, opísanú v prvej kapitole, uvedieme do praxe. Postupným pridávaním ďalších rozmerov do grafov ukážeme, ako zostrojovať prehľadné vizualizácie a dokázať z nich vyčítať potrebné informácie.

Postupne vytvoríme v softvéroch Tableau [5] a R [6] vizualizácie s využitím rôzneho počtu rozmerov. Zostrojíme grafy od dvojrozmerných, cez trojrozmerné a štvorrozmerné, až po päť- alebo viacrozmerné. Na vizualizáciu dvoch dimenzií nám postačí rovina. Pri vizualizácii troch alebo štyroch dimenzií môžeme využiť zobrazovanie informácií v trojrozmernom priestore. Minardova mapa, Obr. 2, v ktorej autor zobrazil až šesť rozmerov, nám potvrdzuje, že viacrozmerné dáta je možné vizualizovať iba s využitím roviny. Avšak, v takom prípade je nevyhnutné nájsť harmonický súlad medzi vizuálnou stránkou a zakomponovaním potrebnej informačnej hodnoty. Vhodnou systematickou vizualizáciou viacrozmerných dát je využitie paralelných súradníc, ktoré budeme v tejto kapitole vytvárať v softvéri Tableau.

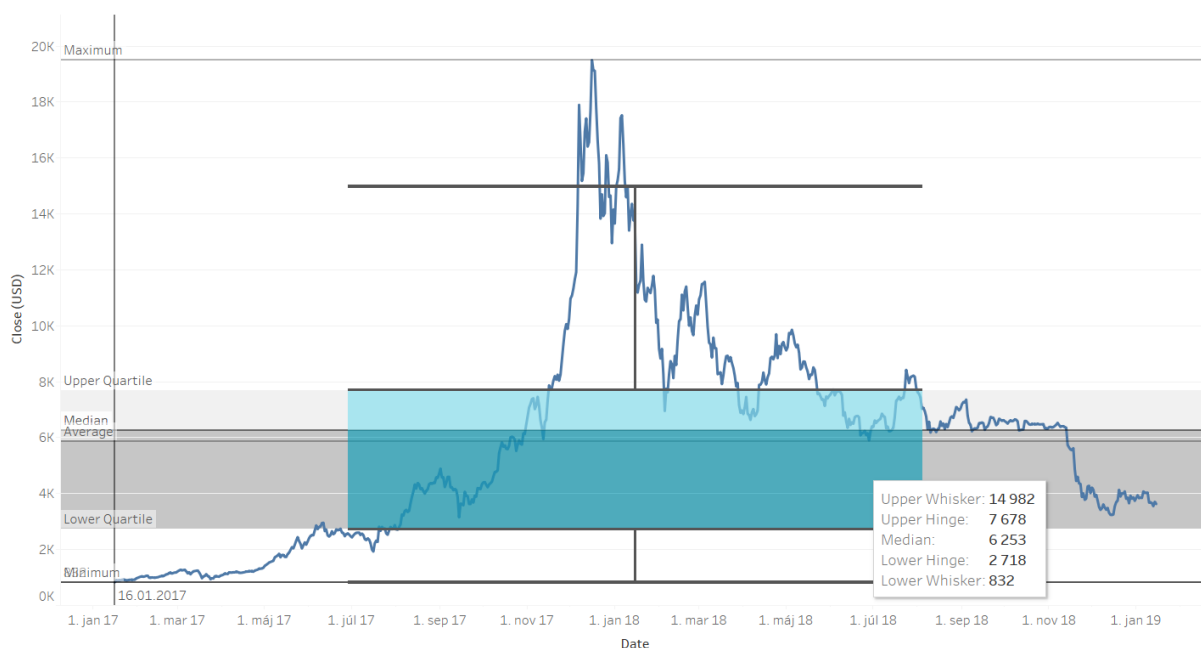
3.1 Čiastočná vizualizácia viacrozmerných dát

Pri akejkoľvek práci s dátami je dôležité im najprv rozumieť. Ak dáta nepoznáme dostatočne dobre, tak výsledné analýzy, štatistiky a vizualizácie nemusia vždy poskytovať požadovaný prínos. Na začiatok sa bližšie zoznámime s jednotlivými premennými v zvolených dátach o Bitcoine, vytvoríme vizualizácie použitím len niektorých rozmerov a získame z nich základné informácie o Bitcoine. Použijeme deskriptívnu štatistiku [15], v ktorej opisujeme dátový súbor pomocou grafov a opisných charakteristík ako miery polohy, variability či tvaru. Následne budeme môcť prejsť aj k vizualizácii a analýze päť- a viacrozmerných dát.

3.1.1 Dvojrozmerné vizualizácie

Na vytvorenie predstavy, ako sa vyvíjala cena kryptomeny Bitcoin v priebehu dvoch rokov slúži nasledujúci graf Obr. 11. Zvolili sme si dva rozmery – časovú zložku sme umiestnili na horizontálnu os a uzatváraciu trhovú cenu Bitcoinu na vertikálnu os. Vďaka kvalitnému

a interaktívnemu používateľskému prostrediu softvéru Tableau je možné získať základné informácie o dátach bez predchádzajúceho programovania.



Obr. 11 - Graf vývoja trhovej ceny Bitcoinu v priebehu dvoch rokov vytvorený v Tableau

Miery polohy charakterizujú dáta na číselnej osi a zaraďujeme sem napríklad strednú hodnotu. Z vizualizácie vytvorenej v Tableau sme zistili, že aritmetický priemer uzatváracej trhovej ceny za sledované obdobie bol 5 848.89 USD. Medián, ktorý je robustnou štatistikou na rozdiel od aritmetického priemeru, bol 6 252.71 USD za sledované obdobie. Robustné štatistiky nie sú citlivé na extrémne hodnoty, a teda poskytujú presnejšiu informáciu o povahe dát.

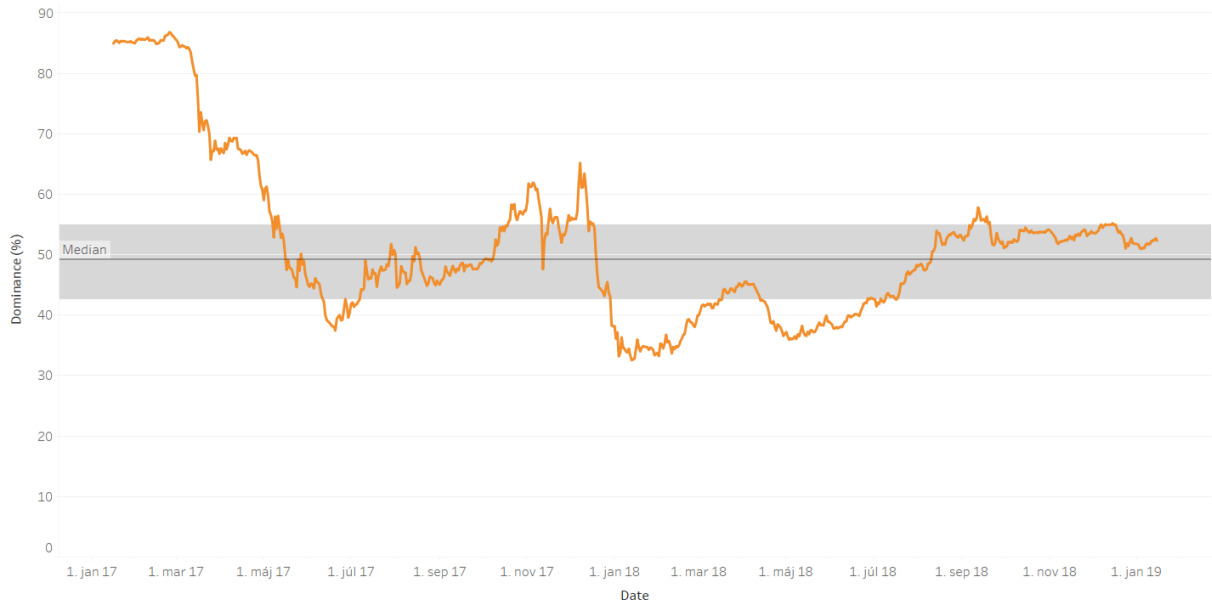
Miery variability hovoria o tom, ako veľmi sa hodnoty v dátovom súbore podobajú, resp. odlišujú. Medzi miery variability patrí variačné rozpätie, medzikvartilové rozpätie, smerodajná odchýlka či rozptyl [15]. V grafe Obr. 11 sme si priamkami zobrazili maximálnu a minimálnu uzatváraciu trhovú cenu za sledované obdobie. Cena dosiahla minimum na začiatku vybraného obdobia, t. j. dňa 16.01.2017 vo výške 831.53 USD. Cena dosiahla maximum dňa 16.12.2017 v hodnote 19 497.40 USD. Variačné rozpätie si vieme jednoducho dopočítať ako rozdiel maxima a minima, ktoré tvorí veľmi vysokú mieru variability, až 18 665.87 USD.

Softvér Tableau umožňuje priamo do grafu zakresliť diagram rozdelenia hodnôt v dátovom súbore na základe kvartilov, tzv. boxplot. Boxplot slúži na predbežnú analýzu dát

a zvyčajne má päť prvkov – dolnú hradbu (z angl. lower whisker), prvý kvartil, medián, tretí kvartil a hornú hradbu (z angl. upper whisker) [20]. Kvartily delia usporiadanú množinu hodnôt na štyri rovnako veľké časti. Teda pod prvým, dolným, kvartilom leží štvrtina hodnôt a nad tretím, horným, kvartilom leží tiež štvrtina hodnôt. Jednotlivé prvky, ktoré sa môžu vyskytnúť v boxplote, sú podrobnejšie vysvetlené v literatúre [20]. Z vizualizácie sme zistili, že prvý kvartil $x_{0.25}$ je na úrovni 2 719 USD a tretí kvartil $x_{0.75}$ je na úrovni 7 672 USD. Dopočítame si rozdiel prvého a tretieho kvartilu a získame informáciu o medzikvartilovom rozpätí, tzv. IQR (z angl. interquartile range), ktoré tvorí 4 953 USD a je robustnou štatistikou. Na Obr. 11 je medzikvartilové rozpätie znázornené dvomi odtieňmi tyrkysovej farby.

Každý softvér sa odlišuje spôsobom naprogramovania. To môže zapríčiniť aj drobné rozdiely v použitých technikách výpočtov a vizualizácií. Pomocou medzikvartilového rozpätia je možné v softvéri R dopočítať hodnotu dolnej, resp. hornej hradby ako $(x_{0.25} - 1.5 \cdot IQR)$, resp. $(x_{0.75} + 1.5 \cdot IQR)$. Zistili sme, že dolná hradba je -4 710.41 USD a horná hradba je 15 101.72 USD. Avšak softvéry Tableau ani R nevykresľujú v boxplote presne tieto hradby. Kolmé línie vychádzajúce zo stredu boxplotu vyjadrujú variabilitu dát pod prvým kvartilom a nad tretím kvartilom, nazývajú sa aj závesy (z angl. whiskers) a obyčajne tvoria 1,5-násobok medzikvartilového rozpätia. Práve tie sú zobrazované v boxplote. Koniec závesov je na oboch stranách určený najnižšou hodnotou dát zo súboru, ktorá nepresahuje dolnú hradbu a najvyššou hodnotou dát zo súboru, ktorá nepresahuje hornú hradbu [20]. Teda zvyčajne je v boxplote označené dolnou hradbou minimum a hornou hradbou maximum, ak neexistujú žiadne odľahlé hodnoty (z angl. outliers). V našom praktickom prípade je hodnota horného závesu 14 982.10 USD a hodnota dolného závesu 831.53 USD.

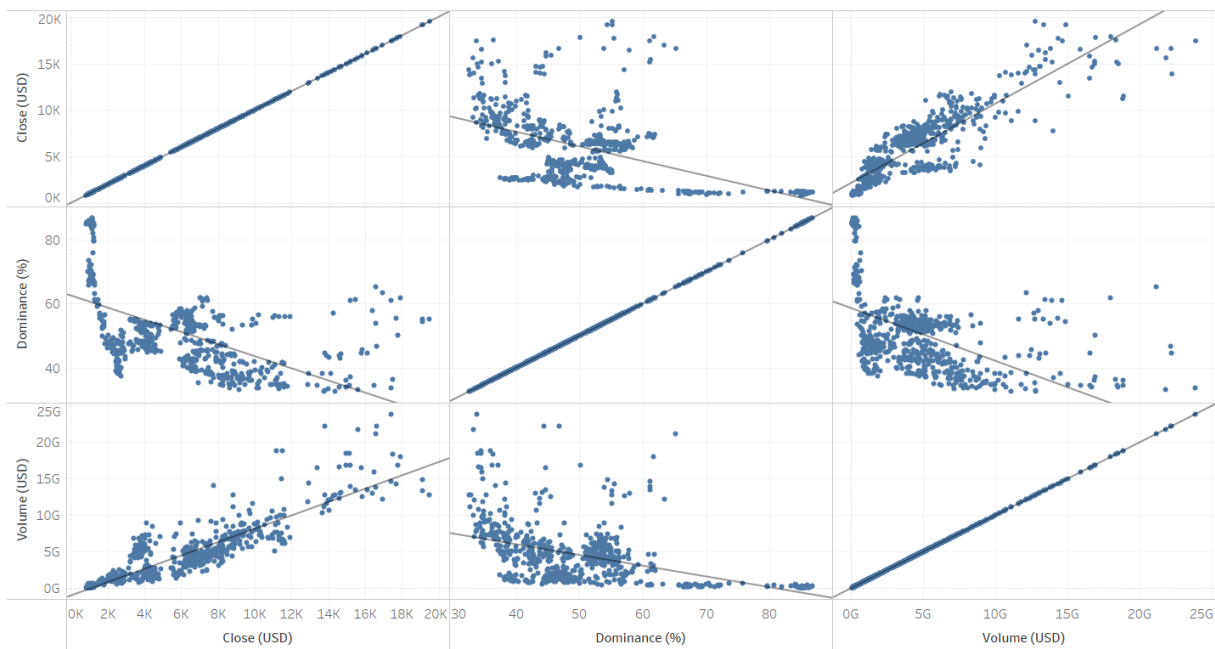
Vďaka boxplotu dokážeme rýchlo identifikovať odľahlé hodnoty, ktoré sa nachádzajú pod dolnou hradbou, resp. závesom alebo nad hornou hradbou, resp. závesom. V softvéri R sú všetky extrémne hodnoty zobrazené v podobe bodov, Obr. 12. Ak o nich chceme zistiť viac informácií, musíme využiť naše znalosti z programovania. V softvéri Tableau máme možnosť priamo z vizualizácie zistiť výskyt odľahlých hodnôt v konkrétnom časovom období a ich presnú hodnotu, Obr. 11.



Obr. 13 - Dvozmerná vizualizácia vývoja dominancie Bitcoinu vytvorená v Tableau

Pozrime sa aj na jednotlivé premenné a ich vzájomný vzťah. Zo štatistiky vieme, že kovariancia vyjadruje závislosť medzi dvomi náhodnými veličinami. Ak je kovariancia nulová medzi dvomi premennými, tak tieto premenné sú nekorelované, a teda nezávisia jedna od druhej. Praktickejším ukazovateľom závislosti medzi dvomi alebo viacerými náhodnými veličinami sa javí korelácia. Jej intenzitu vyjadruje hlavne korelačný koeficient, ktorý predstavuje kovarianciu medzi dvomi normovanými náhodnými veličinami. Hodnoty korelácie sa pohybujú v intervale $< -1; 1 >$, čím je umožnené rýchlejšie identifikovanie miery závislosti premenných.

Spojením viacerých dvozmerných vizualizácií môžeme v Tableau vytvoriť maticu korelácie, Obr. 14. My sme si zvolili na hľadanie vzájomných súvislostí tri premenné z dátového súboru, ktoré má zmysel vizualizovať. Sú nimi uzatváracia trhovacia cena Bitcoinu, jeho percentuálna dominancia spomedzi ďalších kryptomien a objem transakcií uskutočnených v daný trhoví deň.



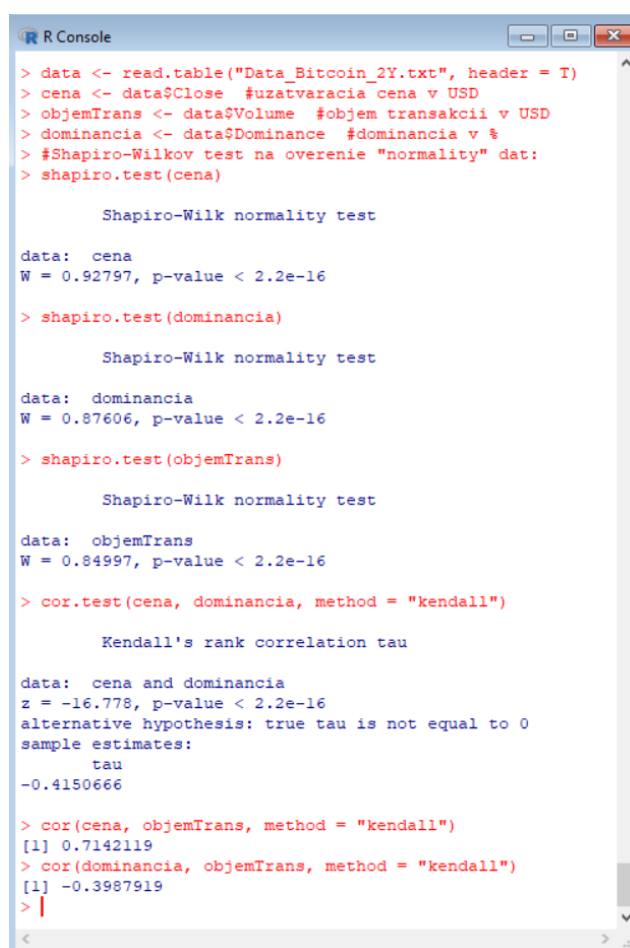
Obr. 14 - Korelačná matica vytvorená v Tableau

Stanovíme si nulovú hypotézu, že naše dáta pochádzajú z normálneho rozdelenia. Oproti nej položíme alternatívnu hypotézu, že dáta nepochádzajú z normálneho rozdelenia. Pravdivosť nulovej hypotézy overíme pomocou Shapiro-Wilkovho testu [17]. Po výpočte v softvéri R sme zistili, že naše dáta nepochádzajú z normálneho rozdelenia, Obr. 15. Keďže hodnota pravdepodobnosti p je vo všetkých troch prípadoch menšia ako hladina významnosti $\alpha = 0.05$, tak hypotézu o tom, že dáta pochádzajú z normálneho rozdelenia, sme museli zamietnuť.

V softvéri R sme si vypočítali koreláciu medzi jednotlivými premennými Kendallovou metódou [21]. Korelácia sa v R štandardne počíta Pearsonovou metódou [22], ktorá vyžaduje normalitu dát. Metódy, ktoré nepožadujú, aby dáta pochádzali z normálneho rozdelenia, sú napríklad Spearmanova [21] a Kendallova. Korelačné koeficienty oboch metód sú robustné štatistiky a je vhodné ich použiť, keď očakávame v štatistickom súbore nejakú odľahlú hodnotu.

Kendallov korelačný koeficient medzi hodnotami uzatváracej trhovej ceny a dominancie je -0.415 , medzi hodnotami uzatváracej trhovej ceny a objemu transakcií je 0.714 , a medzi hodnotami dominancie a objemu transakcií je -0.399 . Kendallov korelačný koeficient hovorí o tom, ako veľmi závisí hodnota jednej premennej od druhej. Záporná hodnota naznačuje nepriamu závislosť, a je zobrazená na Obr. 14 klesajúcim lineárnym trendom, teda priamkou prechádzajúcou pomedzi namerané hodnoty. Môžeme si všimnúť aj podľa výpočtov v R,

programový kód s príslušnými výsledkami je na Obr. 15, aj podľa matice korelácie v Tableau, že závislosť medzi cenou a dominanciou, alebo medzi dominanciou a objemom transakcií nie je veľmi silná. Hodnoty sú výrazne rozptýlené a nedá sa s istotou povedať, že s narastajúcou dominanciou Bitcoinu klesá jeho trhovú cenu či objem uskutočnených transakcií nad Bitcoinom. Avšak, pomerne vysoká korelácia je medzi trhovou cenou a objemom transakcií, ktorú vidíme aj na príslušných dvoch grafoch korelačnej matice Obr. 14. S narastajúcou trhovou cenou Bitcoinu má tendenciu priamoúmerne rásť aj objem vykonaných transakcií. Pre investorov je dôležité sledovať závislosť medzi cenou a objemom transakcií ako aj samotný vývoj objemu transakcií, pretože im to môže pomôcť pri predvídaní budúcich trendov.



```
> data <- read.table("Data_Bitcoin_2Y.txt", header = T)
> cena <- data$Close #uzatvaracia cena v USD
> objemTrans <- data$Volume #objem transakcii v USD
> dominancia <- data$Dominance #dominancia v %
> #Shapiro-Wilkov test na overenie "normality" dat:
> shapiro.test(cena)

      Shapiro-Wilk normality test

data:  cena
W = 0.92797, p-value < 2.2e-16

> shapiro.test(dominancia)

      Shapiro-Wilk normality test

data:  dominancia
W = 0.87606, p-value < 2.2e-16

> shapiro.test(objemTrans)

      Shapiro-Wilk normality test

data:  objemTrans
W = 0.84997, p-value < 2.2e-16

> cor.test(cena, dominancia, method = "kendall")

      Kendall's rank correlation tau

data:  cena and dominancia
z = -16.778, p-value < 2.2e-16
alternative hypothesis: true tau is not equal to 0
sample estimates:
      tau
-0.4150666

> cor(cena, objemTrans, method = "kendall")
[1] 0.7142119
> cor(dominancia, objemTrans, method = "kendall")
[1] -0.3987919
> |
```

Obr. 15 - Programový kód ku korelačnej maticici v R

3.1.2 Trojrozmerné vizualizácie

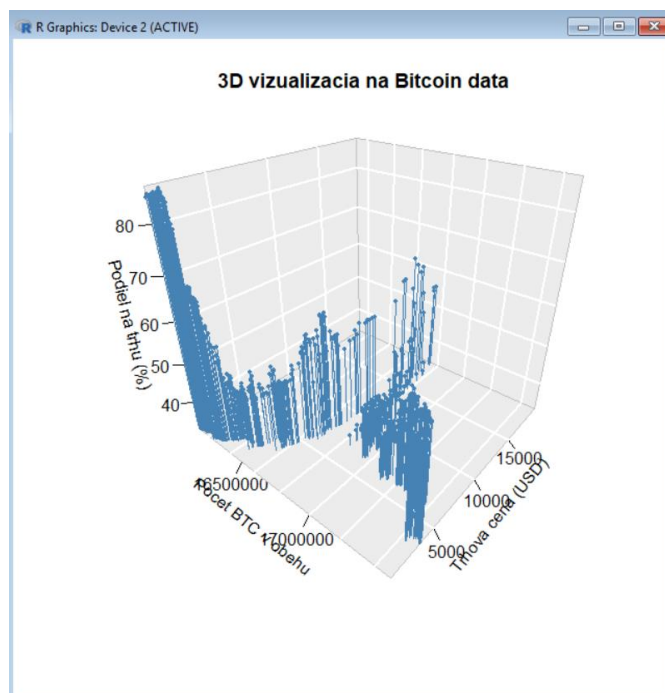
Trojrozmerné dáta zvykneme vizualizovať v priestore, ale taktiež aj v rovine môžu tvoriť prehľadný celok. Príkladom je vizualizácia vytvorená v Tableau, kde na horizontálnej osi sa nachádza časová zložka a na protíahlých vertikálnych osiach sme umiestnili uzatváraciu trhovú cenu Bitcoinu a dominanciu Bitcoinu na trhu v porovnaní s ďalšími kryptomenami.

Informácie, ktoré sme získali v dvojrozmerných grafoch v predchádzajúcej podkapitole, vieme vďaka trojrozmernému grafu Obr. 16 lepšie pochopiť. Keďže sa nám podarilo dosiahnuť, že súčasne vidíme vývoj trhovej ceny aj percentuálneho podielu na trhu, tak vzniká potreba hľadať súvislosti medzi týmito javmi.



Obr. 16 - Porovnanie ceny a dominancie Bitcoinu na časovej osi vytvorené v Tableau

Na ukážku trojrozmernej vizualizácie v priestore sme využili softvér R a funkciu z balíčka na tvorbu trojrozmerných grafov, Obr. 17. Na x -ovú os sme umiestnili informáciu o počte Bitcoinov v obehu, na y -ovej osi sa nachádza uzatváracia trhovú cenu a na osi z je zaznamenaná dominancia Bitcoinu spomedzi ostatných kryptomien. Počet Bitcoinov v obehu môže z určitého pohľadu predstavovať aj časovú zložku vizualizácie, keďže každým dňom narastá počet dobytých Bitcoinov, až do momentu, kým sa nedosiahne maximum stanovené na hodnotu 21 miliónov. Programový kód zo softvéru R je uvedený na Obr. 18.



Obr. 17 - Trojrozmerná vizualizácia dát v priestore vytvorená v R

```

C:\Users\Veronika\Documents\bitcoin analysis.R - R Editor
data <- read.table("Data_Bitcoin_2Y.txt", header = T)
install.packages("plot3D")
library("plot3D")
x <- data$CirculatingSupply
y <- data$Close
z <- data$Dominance
scatter3D(x, y, z, colvar = NULL, col = "steelblue", pch = 19, cex = 0.6,
         bty = "g", phi = 30, type = "h",
         ticktype = "detailed", main = "3D vizualizacia na Bitcoin data",
         xlab = "Pocet BTC v obehu", ylab = "Trhova cena (USD)", zlab = "Podiel na trhu (%)")

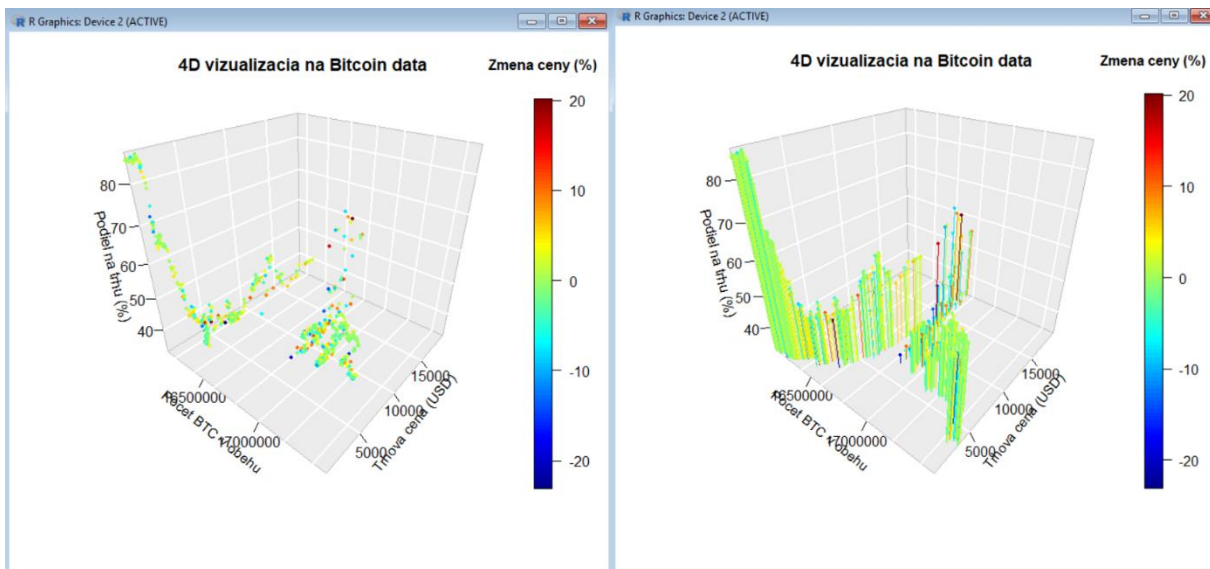
```

Obr. 18 - Programový kód trojrozmernej vizualizácie v R

3.1.3 Štvorrozmerné vizualizácie

Pri vizualizovaní štyroch dimenzií vybraných dát je možné pridať do trojrozmerného priestoru farbu ako ďalší rozmer. My sme sa rozhodli do grafu Obr. 17 pridať farebné rozlíšenie relatívnej zmeny uzatvárajacej trhovej ceny v percentách pre každý zobrazený bod. Farebná škála na Obr. 19 indikuje rozpätie medzi najnižšou a najvyššou percentuálnou zmenou, ktorá počas sledovaného obdobia nastala. Zvislé čiary na pravom grafe slúžia na jednoduchšie identifikovanie polohy bodov v rovine xy . Doplnený programový kód uvádzame na Obr. 20.

Všimnime si, že najvýraznejší nárast, resp. pokles ceny zo dňa na deň sa dial práve v období, keď trhovú cenu Bitcoinu dosahovala svoje maximálne hodnoty. Zmeny boli natoľko veľké, že bolo veľmi ťažké predvídať vývoj v ďalšom trhovom dni. Podobné kolísanie ceny ešte pretrvávalo začiatkom roku 2018, kým sa postupne začalo opäť ustáľovať na nepatrných cenových poklesoch a nárastoch.



Obr. 19 - Štvorrozmerná vizualizácia dát v priestore vytvorená v R

```

C:\Users\Veronika\Documents\bitcoin analysis.R - R Editor
scatter3D(x, y, z, colvar = data$Difference,
  pch = 19, cex = 0.6, bty = "g", phi = 30, type = "h",
  ticktype = "detailed", main = "4D vizualizacia na Bitcoin data",
  xlab = "Pocet BTC v obehu", ylab = "Trhova cena (USD)", zlab = "Podiel na trhu (%)",
  clab = "Zmena ceny (%)")

```

Obr. 20 - Doplnený programový kód štvorrozmernej vizualizácie v R

Približne v treťom kvartáli roku 2017, keď sa počet Bitcoinov v obehu pohyboval okolo čísla 16.5 miliónov, tak trhovú cenu zaznamenávali značné výkyvy. Cena síce podstatne rástla, ale následne aj rýchlo klesala. Napríklad 14. septembra 2017 klesla uzatváracia trhovú cenu o rekordných -23.06 % oproti predchádzajúcemu dňu. S týmito informáciami bola predpoveď ďalšieho vývoja nejasná, a málokomu mohlo napadnúť, že cena Bitcoinu vzrastie šesťnásobne v priebehu len pár mesiacov. Otázkou ostáva, či sa bude podobná anomália v prípade Bitcoinu niekedy opakovať. Odpoveď na ňu sa môžeme pokúsiť nájsť prostredníctvom vizualizácie vytvorenej metódou paralelných súradníc.

3.2 Tvorba paralelnej súradnicovej sústavy

Vizualizácia päť- a viacrozmerných dát si vyžaduje použitie zložitejšej metódy. My sme sa rozhodli použiť metódu paralelných súradníc [2] a pomocou nej vizualizovať dáta o kryptomene Bitcoin v softvéri Tableau. Tento softvér neponúka vytvorenie paralelnej súradnicovej sústavy v základnom balíčku funkcionalít, preto podobná vizualizácia je časovo a programovo náročnejšia. V tejto kapitole opíšeme presný postup vytvárania vizualizácie metódou paralelných súradníc, aby si čitateľ dokázal vytvoriť obdobnú vizualizáciu na analýzu vlastných dát. Vizualnú pomôcku k zorientovaniu sa v texte poskytuje Príloha B, v ktorej je zdokumentovaný celý postup v obrázkoch.

V prvom rade si vytvoríme v Tableau novú záložku (z angl. Worksheet). V ľavej časti obrazovky sa nachádzajú jednotlivé stĺpce zvoleného dátového súboru, ktoré sú rozdelené na dimenzionálne parametre (z angl. Dimensions) a parametre meraní (z angl. Measures), príklad obrazovky môžeme vidieť na priloženom obrázku Obr. B1. Pred tvorbou grafu by sme si mali určiť, ktoré stĺpce z dátového súboru budeme v paralelných súradniciach zobrazovať. V prípade našich dát chceme zobrazovať v paralelných súradniciach dátum, počet Bitcoinov v obehu, objem transakcií nad Bitcoinom, uzatváraciu trhovú cenu a percentuálnu dominanciu spomedzi ďalších kryptomien. Vybrané merania je potrebné zobrazovať na paralelných osiach s rovnakým škálovaním. Z tohto dôvodu musíme každý z vybraných parametrov znormalizovať, aby sme ho vedeli zobraziť na škále (0;1). Na začiatok budeme pracovať s parametrami v časti Measures, neskôr doplníme do vizualizácie dátumovú zložku predstavenú v časti Dimensions.

Zvolíme si prvý parameter z časti Measures, s ktorým chceme pracovať. Po kliknutí pravým tlačidlom sa otvorí menu, v ktorom vyberieme možnosť duplikácie meradla (z angl. Duplicate), Obr. B2. Vytvorí sa kópia parametra, ktorú je možné po kliknutí pravým tlačidlom upravovať (z angl. Edit), Obr. B3. Môžeme zmeniť názov duplikovaného parametra, a taktiež upravíme jeho výpočet. Pokiaľ sa nejedná o agregované hodnoty, tak pre normalizáciu parametra použijeme vzorec

$$\frac{(MIN([Measure]) - TOTAL(MIN([Measure])))}{(TOTAL(MAX([Measure])) - TOTAL(MIN([Measure])))}$$

kde [*Measure*] predstavuje názov parametra merania, ktorý normalizujeme na os (0; 1). Postup normalizácie zopakujeme pre všetky ostatné parametre meraní, ktoré budeme chcieť vizualizovať metódou paralelných súradníc, Obr. B4.

Ďalším krokom je vytvorenie prvého konceptu vizualizácie. Z časti Dimensions vezmeme položku „Measure Names“ a presunieme ju do lišty „Columns“, aby sme názvy parametrov zobrazovali na grafe v stĺpcoch. Z časti Measures vezmeme položku „Measure Values“ a presunieme ju do lišty „Rows“, aby jednotlivé hodnoty parametrov boli zobrazované v riadkoch. Týmto spôsobom sme vytvorili sústavu s názvami parametrov na horizontálnej osi a hodnotami parametrov na vertikálnej osi. Po kliknutí pravým tlačidlom myši na „Measure Names“ alebo „Measure Values“ sa otvorí menu, kde zvolíme zobrazenie filtra na obrazovke (z angl. Show Filter), Obr. B5. Vo filtri vyberieme iba tie merania, ktoré sme manuálne normalizovali, Obr. B6.

Na vykreslenie lomených čiar je nevyhnutné zvoliť si dimenziu, v ktorej chceme jednotlivé hodnoty zobrazovať. V prípade našich dát je to dátum, ale mohla by to byť aj iná dimenzia, napríklad identifikačné číslo záznamu. Neskôr uvidíme v ďalšom postupe, že je potrebné dodržiavať konzistentnosť po zvolení dimenzie, pretože na jej základe budeme vypočítavať aj normalizované merania. Vybranú dimenziu – dátum – prenesieme do časti „Marks“, Obr. B7, a zmeníme typ zobrazenia na priamku (z angl. Line), Obr. B8. Získali sme súbor jednofarebných lomených čiar, ktoré pretínajú paralelné osi na škále (0; 1). Pre lepšiu orientáciu vo vizualizácii je vhodné zvoliť farebné rozlíšenie na základe jedného z parametrov meraní. Záleží na používateľovi, aký rozmer si vyberie na zobrazovanie prostredníctvom farebnej škály. My sme sa na začiatok rozhodli pre farebné rozlíšenie na základe presného dátumu, a teda dátumovú zložku z časti Dimensions sme potiahli na ikonu „Color“ v časti „Marks“, Obr. B9. Upravili sme farebnú škálu aj priesvitnosť lomených čiar podľa našich potrieb, Obr. B10.

Do vytvoreného grafu s lomenými čiarami pridáme osi, aby mohla vzniknúť paralelná súradnicová sústava. Opäť zoberieme z časti Measures meranie „Measure Values“ a pridáme ju druhýkrát na lištu „Rows“, aby sme ju mohli v časti „Marks“ samostatne upravovať, Obr. B11. V Dimensions zoberieme dátumovú zložku a potiahneme ju na ikonu „Path“ v časti „Marks“. Typ zobrazenia zmeníme na priamku „Line“ a vďaka tomu vytvoríme rovnobežné úsečky

predstavujúce osi jednotlivých meraní, Obr. B12. V tomto momente máme vytvorené lomené čiary a paralelné osi zobrazené samostatne na pracovnej ploche.

Vizualizácie v Tableau sú interaktívne a pri ich analyzovaní si môže používateľ posúvaním myšky zobrazovať prehľad informácií na obrazovke. Informácie na informačnom štítku sa dajú jednoducho upravovať cez ikonu „Tooltip“ v časti „Marks“, Obr. B13. V našej vizualizácii nechceme zobrazovať informácie, ak sa posúvame po jednej z paralelných osí, preto v okne na úpravu informačného štítku (z angl. Tooltip) pre parametre meraní znázorňujúce osi „Measure Values (2)“ vymažeme existujúci text. Poznámky s informáciami o jednotlivých meraniach budeme chcieť zobrazovať pri prechádzaní myšou po lomených čiarach. V meraniach znázorňujúcich lomené čiary, t. j. „Measure Values“, musíme do časti „Marks“ pridať parametre s hodnotami, ktoré budeme chcieť zobrazit' cez „Tooltip“. Tieto parametre, ktoré predstavujú pôvodné a nie normalizované hodnoty dátového súboru, pridáme potiahnutím na ikonu „Detail“. Následne cez ikonu „Tooltip“ môžeme podľa potreby upravovať text na informačnom štítku, a zároveň zobrazovať hodnoty, ktoré sa nachádzajú v meraniach ako „Detail“. Text, ktorý závisí od hodnoty parametra sa mení vzhľadom na vybranú lomenú čiaru vo vizualizácii, Obr. B14.

Prehľadnosť grafu môžu zvýšiť popisy jednotlivých osí. V časti „Marks“ si vyberieme parametre meraní, ktoré zobrazujú paralelné osi „Measure Names (2)“ a z časti Dimensions potiahneme „Measure Names“ na ikonu „Label“ v časti „Marks“. Po kliknutí na ikonu „Label“ sa otvorí menu, v ktorom nastavíme „Min/Max“ v časti „Marks to Label“, aby sa názov osi zobrazoval na oboch koncoch úsečky. V časti „Scope“ vyberieme možnosť „Line/Pie“. Týmto spôsobom sme každej osi priradili pomenovanie podľa parametra, ktorý má predstavovať, Obr. B15. Ako bonus môžeme na každú os doplniť informáciu o extrémoch, ktoré boli v danom meraní nadobudnuté. Vytvoríme ďalšie duplikáty pôvodných parametrov a v každej novej premennej zmeníme výpočtový vzorec na podmienku

$$IF [Normalised Measure] = 1 THEN TOTAL(MAX([Measure]))$$
$$ELSEIF [Normalised Measure] = 0 THEN TOTAL(MIN([Measure])) END,$$

kde $[Normalised Measure]$ je meranie, ktorého výpočet sme normalizovali na hodnoty (0; 1) a $[Measure]$ je pôvodná hodnota merania v dátovom súbore, Obr. B16. Názvy nových

parametrov môžeme premenovať, aby sme vedeli neskôr identifikovať v rozrastajúcom sa zozname Measures, že sa jedná o popisy k osiam.

Urobíme ešte niekoľko úprav vo vytvorenej vizualizácii, ktoré nám môžu neskôr uľahčiť situáciu pri uskutočňovaní ďalších zmien v grafe. Pre parametre meraní definujúce lomené čiary upravíme v časti „Marks“ ich vlastnosť – každý parameter, ktorý využívame na určovanie konkrétnej hodnoty na príslušnej osi zmeníme cez pravé kliknutie myši na atribút (z angl. Attribute), Obr. B17. Vlastnosť atribútu zabezpečí, že sa nám vráti vo výsledku práve jedna hodnota zo záznamu v danom stĺpci. V samostatnej časti „Measure Values“ zmeníme spôsob výpočtu normalizovaných parametrov na výpočet pomocou dátumu, t. j. „Compute Using / Date“, resp. zvolenej dimenzie na začiatku vytvárania paralelných súradníc, Obr. B18. Po kliknutí na „Measure Names“ v časti Dimensions sa zobrazí menu, Obr. B19, cez ktoré otvoríme tabuľku s pomenovaniami parametrov (z angl. Aliases). Názvy parametrov, ktoré sa zobrazujú ako názvy paralelných osí, pomenujeme výstižnejšie, Obr. B20.

K premenným definujúcim paralelné osi doplníme cez ikonu „Label“ parametre, ktoré sme si pripravili na zobrazovanie minima a maxima na príslušnej osi. Výpočet týchto popisov nastavíme tiež cez dátumy, t. j. „Compute Using / Date“, resp. inú zvolenú dimenziu, Obr. B21. Po kliknutí na ikonu „Label“ sa otvorí menu úprav a vieme sa dostať do formátovacieho okna pre zobrazujúci sa text, Obr. B22. V tomto prípade môže nastať problém, ak v jednom zázname dátového súboru bolo súčasne nadobudnutých viacero extrémov pre rôzne parametre merania. Extrémne hodnoty sa budú zobrazovať duplicitne na príslušných osiach. Podobná situácia nastala aj s použitím našich dát, ktorú sme ale neskôr vyriešili vytvorením samostatných pracovných hárkov na výpočet minimálnych a maximálnych hodnôt. Výsledné výpočty spolu s vizualizáciou pomocou metódy paralelných súradníc sme spojili do spoločnej tabule (z angl. Dashboard) a vznikla finálna vizualizácia Obr. 21.

Z rozpracovaného grafu odstránime predvolenú hlavičku súradnicových osí, kliknutím pravým tlačidlom na hodnoty na horizontálnej osi alebo na jednej z vertikálnych osí, prípadne na jeden z parametrov „Measure Names“ umiestnených v kolónke „Rows“. V zobrazenom menu odklikneme možnosť „Show Header“, aby sa nezobrazovali automaticky vytvorené osi, Obr. B23. Taktiež odstránime vodiace čiary z grafu. Pravým tlačidlom klikneme na vizualizáciu, v menu zvolíme možnosť „Format“, Obr. B24. Zobrazí sa lišta, v ktorej môžeme meniť rôzne

nastavenia grafu. Vyberieme ikonu „Lines“ a na záložke „Rows“ v riadku „Grid Lines“ zvolíme možnosť „None“, Obr. B25. Týmto spôsobom sa nám podarilo z grafu odstrániť nepotrebné elementy a sprehľadniť doteraz vytvorenú vizualizáciu.

Posledným krokom vo vytvorení paralelnej súradnicovej sústavy je spojenie grafu lomených čiar a grafu paralelných osí. V softvéri Tableau existuje možnosť zobrazenia dvoch grafov v jednej súradnicovej sústave. Po kliknutí pravým tlačidlom myši na jedno z meradiel v riadku „Rows“ si v menu vyberieme možnosť „Dual Axis“, Obr. B26. Vertikálne súradnicové osi sa zobrazia v jednej rovine. Aby sme si mohli byť istí, že sa osi nachádzajú na rovnakej úrovni, tak sa odporúča urobiť synchronizáciu osí (z angl. Synchronize Axis). Túto možnosť nájdeme v menu po kliknutí pravým tlačidlom na „Header“, ktorý musíme na chvíľu opätovne zobraziť. Po zosúladení oboch osí je vizualizácia dát metódou paralelných súradníc hotová a výsledok je zobrazený na Obr. B27.

Dôležitou úlohou bolo vytvoriť v Tableau funkčnú vizualizáciu. Splnením tejto podmienky môžeme dopĺňať do paralelných súradníc ďalšie merania podľa vyššie popísaného postupu, meniť ich poradie, upravovať finálny vzhlľad vizualizácie, či analyzovať dáta a hľadať medzi nimi skryté súvislosti. My sme doteraz pracovali pri vytváraní paralelných súradníc iba so štyrmi premennými – počet Bitcoinov v obeh, uzatváracia trhovacia cena, dominancia a objem transakcií. Doplníme do grafu ešte piaty rozmer, ktorým bude dátumová zložka. V časti Dimensions duplikujeme parameter s dátumami. Upravíme výpočet nového parametra cez možnosť „Edit“. Vzorec upravíme na $ATTR([Measure])$, kde $[Measure]$ predstavuje dátum a funkcia zmeny premennej na atribút nám zabezpečí vrátenie práve jedného dátumu ku každému záznamu, Obr. B28. Ak sa nová mierka nepresunula automaticky do časti Measures, tak to urobíme manuálne potiahnutím mierky z jedného sektora do druhého.

Vytvoríme duplikát novej mierky, aby sme mohli jej hodnoty normalizovať na os (0; 1). V danom prípade sa jedná o agregovanú hodnotu, preto sa bude vzorec trochu líšiť od toho predchádzajúceho pre neagregované hodnoty, Obr. B29. Vzorec má tvar

$$\frac{([Measure] - WINDOW_MIN([Measure]))}{(WINDOW_MAX([Measure]) - WINDOW_MIN([Measure]))}$$

Novovzniknutú normalizovanú hodnotu dátumovej zložky pridáme do sektora „Measure Values“ k ostatným normalizovaným hodnotám, ktoré využívame vo vizualizácii paralelných súradníc. V možnostiach k normalizovanému dátumu doplníme spôsob výpočtu s využitím dátumovej dimenzie, t. j. „Compute Using / Date“. Vo vizualizácii pribudla nová os a lomené čiary sa predĺžili o daný rozmer, Obr. B30.

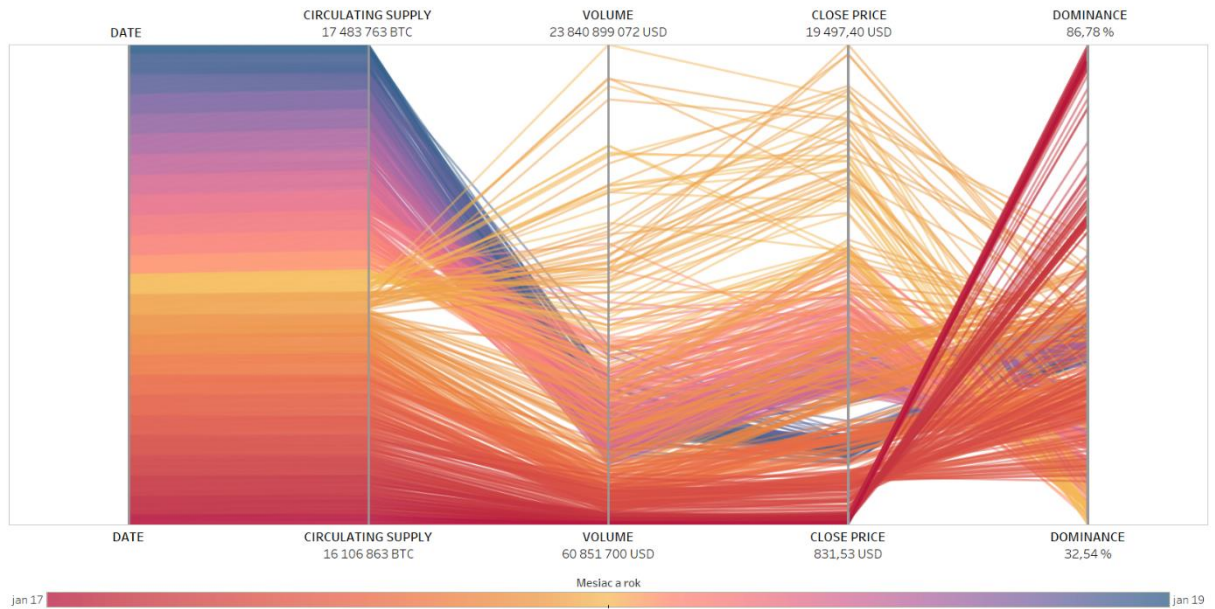
Rovnako, ako predchádzajúce merania, tak aj dátumový parameter premenujeme cez možnosť „Aliases“ na výstižný názov, ktorý sa bude zobrazovať pri osi. V sektore „Measure Values“ presunieme dátumovú os na prvé miesto. Informačnú lištu „Tooltip“ pre lomené čiary môžeme podľa potreby tiež upraviť.

Opäť vytvoríme duplikát dátumového atribútu. Tentokrát bude slúžiť na identifikovanie minimálnej a maximálnej hodnoty v dátumovom stĺpci, Obr. B31. Keďže momentálne pracujeme s agregovanou hodnotou, tak použijeme vzorec

$$IF [Normalised Measure] = 1 THEN WINDOW_MAX([Measure])$$
$$ELSEIF [Normalised Measure] = 0 THEN WINDOW_MIN([Measure]) END.$$

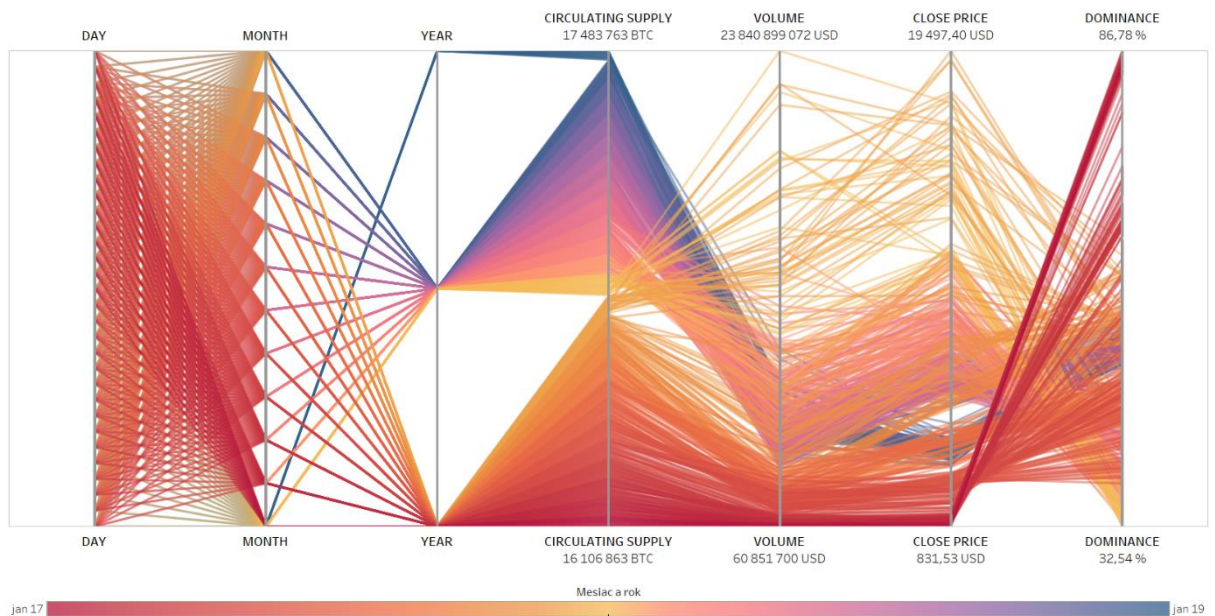
Vytvorený parameter s podmienkou umiestnime cez ikonu „Label“ k ostatným popisom osí. Vo formátovacom okne pre text popisov môžeme upraviť jeho zobrazovanie a zjednotiť s existujúcimi, Obr. B32. Avšak, aj pri zobrazovaní najmenšieho a najväčšieho dátumu z dátového súboru môže pretrvávajúť problém so zobrazovaním duplicitných hodnôt. Finálna úprava závisí od typu použitých dát a kreativity používateľa softvéru Tableau.

Náš výsledok z tvorby paralelnej súradnicovej sústavy prezentujeme vizualizáciou s usporiadanými paralelnými osami Obr. 21. Farebnú škálu sme si zvolili podľa mesiacov v danom roku merania a poradie osí sme sa snažili zvoliť takým spôsobom, aby sa dali z grafu nenáročne čítať súvislosti medzi parametrami meraní.



Obr. 21 - Vizualizácia viacrozmerných dát metódou paralelných súradníc v Tableau

Pre účely niektorých analýz je vhodné zobrazovať dni, týždne, mesiace či roky na samostatných osiach. Preto sme vytvorili pre ilustráciu ďalšiu alternatívu vizualizácie metódou paralelných súradníc. V grafe Obr. 22 rozdelíme dátumový rozmer na tri premenné – deň, mesiac a rok.



Obr. 22 - Rozdelená dátumová zložka vo vizualizácii metódou paralelných súradníc v Tableau

4 Vyhodnotenie praktickej časti

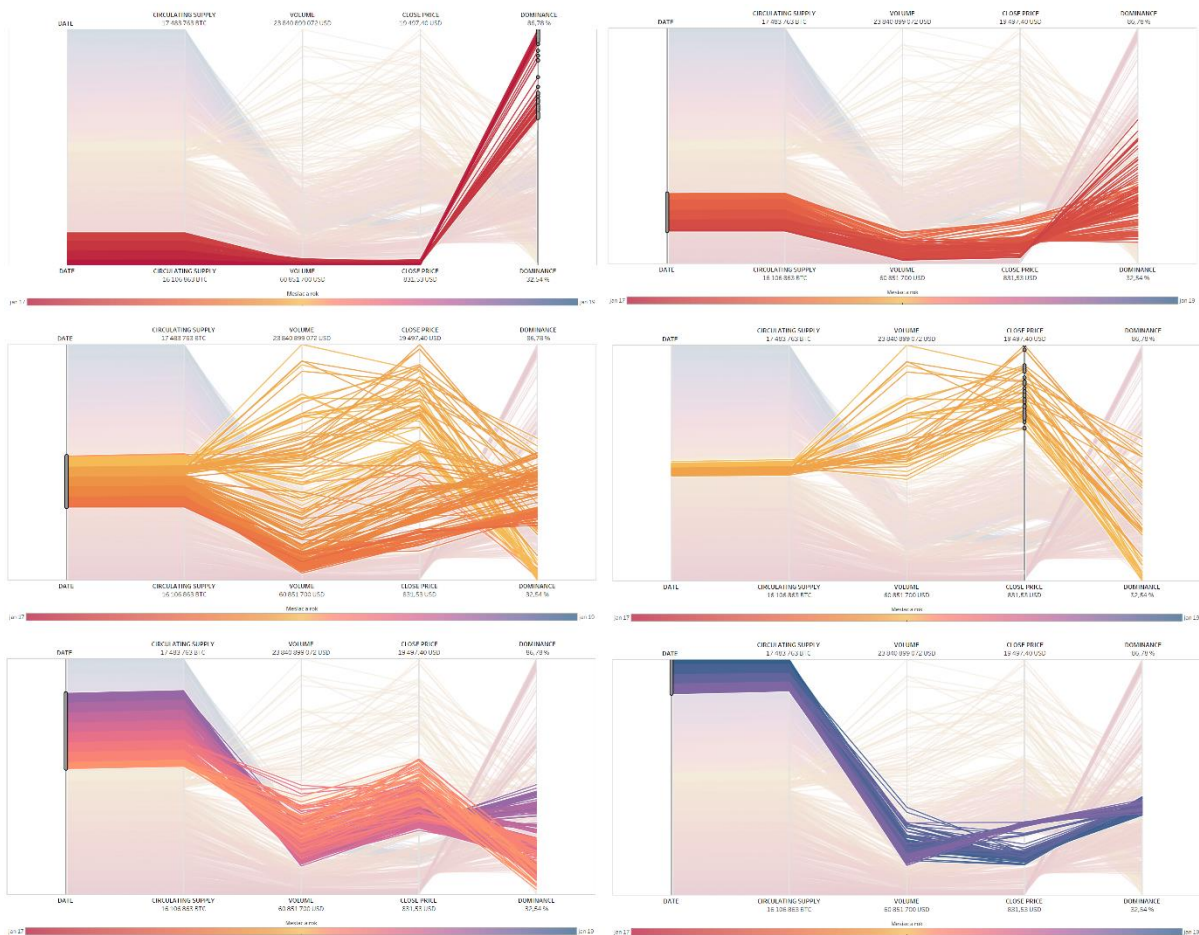
Vizualizácia informácií je viackrokový proces. Na začiatku potrebujeme vedieť, akým spôsobom budeme vizualizovať dostupné dáta. Výber správnej techniky závisí od viacerých faktorov, napríklad od typu dát, ktorými disponujeme, od cieľového publika, ktorému informácie prezentujeme, od spôsobu, akým chceme vysvetliť dáta a súvislosti medzi nimi. Po zvolení vhodnej vizualizačnej metódy môžeme pristúpiť k samotnej vizualizácii dát. Avšak, vytvorením výsledného grafu proces vizualizácie nekončí. Veľmi dôležitým krokom v celej postupnosti je interpretácia vytvoreného obrazu.

V predchádzajúcej kapitole sme uviedli viacero príkladov vizualizácií dát pre rôzny počet zobrazovaných rozmerov, ktorých obsah sme sa snažili čitateľovi ihneď objasniť. Avšak metóda paralelných súradníc je natoľko obsiahla, že podrobnej analýze a vysvetleniu jej prvkov so všetkými vzájomnými súvislosťami musíme udeliť osobitnú pozornosť. V nasledujúcich podkapitolách načrtneme možné vysvetlenia, týkajúce sa najmä metódy paralelných súradníc. Naším cieľom v tejto časti diplomovej práce je priblížiť čitateľovi analytickú stránku grafov, nevyhnutnú súčasť tvorby vizualizácií. Zároveň doplníme informácie, ktoré doteraz neboli dostatočne vysvetlené na praktickom príklade.

4.1 Interpretácia vizualizácie metódou paralelných súradníc

Pre odvetvia, ktoré pracujú na dennej báze s enormnými objemami dát, môže byť ich zobrazenie na dvojrozmernej ploche veľkým prínosom k ďalším výsledkom a výskumom. Metóda paralelných súradníc umožňuje objavovať vzťahy medzi premennými viacrozmerných dát. Významnou súčasťou interpretácie informácií je stanovenie si cieľa, čo chceme z vizualizácie zistiť. V prípade dát o kryptomene Bitcoin by nás mohlo zaujímať, aké sú skryté zákonitosti medzi jednotlivými parametrami, ktoré by mohli ovplyvniť rozhodnutia investora pri obchodovaní s touto kryptomenou.

Existuje viacero možností, akými sa dajú dáta z vizualizácie interpretovať. Jedným zo základných pravidiel patrí skúmanie javov, ktoré v grafe upútali našu pozornosť. Vďaka interaktívnemu prostrediu softvéru Tableau sme si vyznačili šesť prípadov, ktoré môže mať zmysel podrobnejšie skúmať. Ukážky vybraných výsledkov sme zhrnuli na Obr. 23.



Obr. 23 - Ukážky z analýzy vizualizácie informácií metódou paralelných súradníc v Tableau

Očividným faktom je, že medzi trhovým dňom a počtom Bitcoinov v obehu je priama závislosť – zobrazujú to lomené čiary medzi prvou a druhou paralelnou osou. Pokiaľ z kryptomeny Bitcoin nebolo vyťažené maximum, stanovené na dvadsaťjeden miliónov kusov, tak sa do obehu bude každým dňom dostávať väčší počet tejto kryptomeny, a teda budú tieto dve hodnoty priamoúmerne rásť.

Na základe prvých dvoch ukážok z dátovej analýzy Obr. 23 vidíme, že v prvej polovici roku 2017 bola relatívna stagnácia, prípadne veľmi pomalý rast objemu transakcií vykonaných nad Bitcoinom a jeho trhovej ceny. Napriek tomu stále klesala dominancia Bitcoinu. Mohlo to byť spôsobené vstupom ďalších kryptomien na trh a ich narastajúca popularita. V súčasnosti je na stránke CoinMarketCap [3] evidovaných viac ako dvetisíc rôznych kryptomien. Taktiež mohli investori začať v tom období odpredávať svoje podiely z dôvodu jemne zvýšenej trhovej ceny a investovať do iných kryptomien, čím dominancia Bitcoinu klesla.

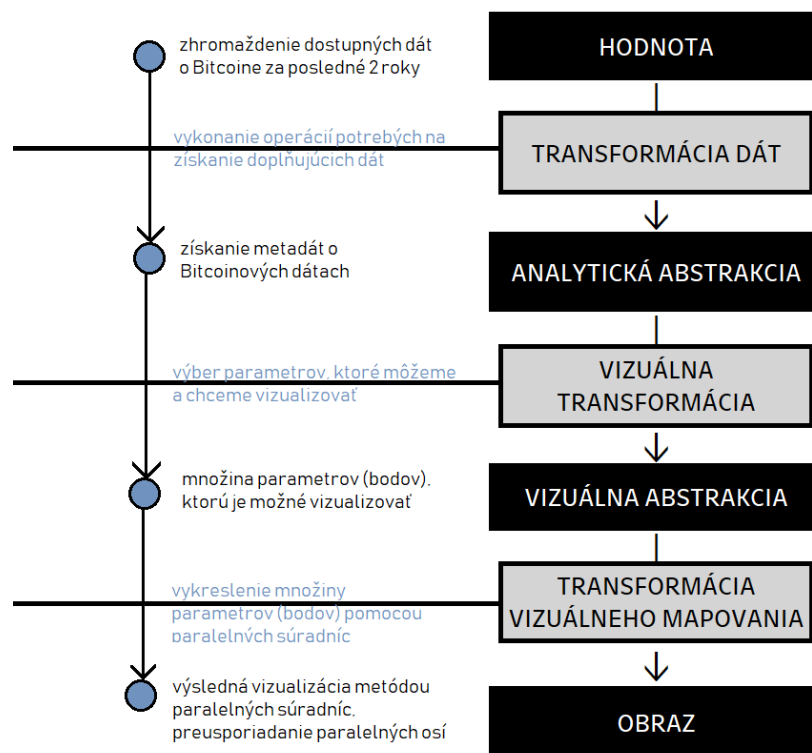
Na prelome rokov 2017 a 2018 bola situácia na Bitcoinovom kryptotrhu veľmi ťažko predvídateľná. Podľa oranžovožltých lomených čiar vidíme, že práve toto obdobie sa vyznačovalo výrazným nárastom trhovej ceny a trvalo približne pol roka – trhovú cenu je definovaná na štvrtej paralelnej osi. Objem transakcií, cena aj dominancia najprv priamoúmerne narastali, ľudia obchodovali čoraz viac s Bitcoinom. Lenže, keď sa cena vyšplhala približne na úroveň deviatich tisíc amerických dolárov za jeden Bitcoin, tak popularita Bitcoinu rýchlo klesla. Konzervatívni investori sa začali zbavovať svojich podielov. S nepredvídateľne rastúcou cenou sa objem transakcií zvyšoval veľmi opatrne, no výrazným trendom bola stále upadajúca dominancia.

Predposledná ukážka z analýzy vizualizácie hovorí o akomsi ustáľovaní sa situácie na trhu s kryptomenou Bitcoin. Podľa zhluku križujúcich sa čiar medzi štvrtou a piatou paralelnou osou by sme mohli hovoriť o náznaku negatívnej korelácie medzi dominanciou a cenou. Čím bola trhovú cena vyššia v roku 2018, tým nižšiu dominanciu kryptomeny sme zaznamenávali. Príkladom pozitívnej korelácie takmer počas celého sledovaného obdobia by mohol byť vzťah objemu transakcií a trhovej ceny. V korelačnej matici Obr. 14 sme ukázali relatívne vysokú koreláciu medzi týmito dvoma parametrami, a v paralelných súradniciach si môžeme túto skutočnosť ešte raz potvrdiť.

Záverečná ukážka s fialovými lomenými čiarami indikuje situáciu na trhu na konci roku 2018 a na začiatku roku 2019. Tu môžeme rýchlo spozorovať výraznú podobnosť s dátami z druhého štvrťroka 2017. Objemy transakcií a trhovú cenu sú relatívne malé, zatiaľ čo dominancia Bitcoinu sa ustáľuje na úrovni päťdesiatich percent. Otázkou pre čitateľa na rozmyšľanie a podrobnejšiu analýzu ostáva, či sa môže podobná situácia s nestabilnými cenami zopakovať a čím konkrétne by mohla byť ovplyvnená.

4.2 Aplikovanie referenčného modelu stavu informácií

Na vizualizáciu viacrozmerých dát o Bitcoine sme použili metódu paralelných súradníc. Pre lepšie pochopenie fungovania tejto techniky ju vieme opísať pomocou referenčného modelu stavu informácií. Grafické vysvetlenie postupu je zobrazené na Obr. 24.



Obr. 24 - Aplikácia referenčného modelu stavu informácií na dáta o Bitcoinu

Na hodnotovej úrovni sa nachádzajú dostupné dáta o Bitcoinu za posledné dva roky. Údaje sme zhromaždili z webových stránok spoločností CoinMarketCap [3] a Blockchain [4], a následne sme ich zaznamenali v tabuľke MS Excel. Transformácia dát nám zabezpečila získanie doplňujúcich informácií o zhromaždených dátach. My sme namiesto manuálneho počítania pomocou vzorcov použili naprogramované funkcie softvérov Tableau [5] a R [6], aby sme získali informácie o strednej hodnote uzatváracj trhovej ceny alebo dominancii Bitcoinu, či údaje o korelácii medzi jednotlivými premennými. V tejto fáze transformácie môžeme zisťovať aj počet záznamov v dátovom súbore a nachádzať prípadné chýbajúce hodnoty. Všetky dodatočne získané informácie o pôvodných, surových dátach tvoria metadáta a umiestňujeme ich na úroveň analytickej abstrakcie v referenčnom modeli.

Akonáhle poznáme náš dátový súbor trochu lepšie a vieme, čo obsahuje, tak môžeme pristúpiť k vizuálnej transformácii. Vyberieme a prípadne aj upravíme na potrebný tvar tie ukazovatele, ktoré sa dajú vizualizovať. My sme si zvolili na vizualizáciu dátumovú zložku, uzatváraciu trhovú cenu Bitcoinu, počet Bitcoinov v obehu, objem transakcií v daný trhovú deň

a jeho dominanciu spomedzi ďalších obchodovaných kryptomien. Tieto parametre tvoria na úrovni vizuálnej abstrakcie množinu bodov, ktorú budeme zobrazovať na grafe.

Poslednou transformáciou v referenčnom modeli je vizuálne mapovanie. Vybranú množinu bodov postupne mapujeme metódou paralelných súradníc na obrazovú úroveň. Výsledný obraz tvorí v našom prípade vizualizácia Obr. 21 vytvorená v Tableau. Na záver môžeme zmeniť usporiadanie paralelných osí či doplniť chýbajúce informácie do informačného štítka, aby sme si uľahčili analýzu informácií.

V tejto záverečnej kapitole sme si zhrnuli najdôležitejšie poznatky, ktoré vyplynuli z vizualizácie zostrojenej pomocou metódy paralelných súradníc. Vysvetlili sme si spôsob, akým sa dajú interpretovať získané informácie z grafu, a taktiež sme uplatnili teóriu referenčného modelu na konkrétnu metódu, s ktorou sme pracovali.

Záver

Cieľom diplomovej práce bolo vytvoriť výučbový materiál, ktorý priblíži tému vizualizácie informácií študentom predmetu Počítačová grafika pre manažérov, ale aj ďalším čitateľom, ktorí sa zaujímajú o spôsoby grafickej reprezentácie dát. Na úvod sme stručne predstavili históriu vizualizácií, v ktorej sme uviedli príklad prepracovanej vizualizácie – Minardovu mapu Napoleonovho ťaženia na Rusko. Mapa, ktorá vznikla v druhej polovici devätnásteho storočia bola schopná veľmi prehľadne zaznamenať až šesť dátových rozmerov na dvojrozmernom plátne. Tento grafický skvost nás neskôr inšpiroval k hľadaniu vizualizačnej metódy, ktorá dokáže znázorniť viacrozmerné dáta bez straty informácie či čitateľnosti.

Dôležitú úlohu pri písaní tejto práce zohralo aj dielo Vaganyana [1]. Vďaka jeho knihe sme čitateľovi ponúkli na výber sedem druhov vizualizácií, ktoré sa najčastejšie využívajú v manažmente. K jednotlivým obrázkom sme vytvorili novodobé ekvivalenty, a tak poskytli náhľad na rozličné spôsoby uplatnenia vizualizácií, nápady na grafické utriedenie dát, aj príklady niekoľkých vizualizačných softvérov. Táto kapitola s porovnaním viacerých druhov grafov odštartovala vznik praktickej časti v diplomovej práci. Vybrali sme si dáta z finančnej oblasti, konkrétne z rozrastajúceho sa trendu kryptomien. Vďaka spoločnostiam CoinMarketCap [3] a Blockchain [4] sme zhromaždili v jednom dátovom súbore dostupné údaje o kryptomene Bitcoin.

Vytvorili sme niekoľko dvojrozmerných, trojrozmerných aj štvorrozmerných vizualizácií s použitím nástrojov Tableau [5] a R [6]. Pri každom vzniknutom grafe sme sa pokúsili interpretovať získané informácie. Vďaka boxplotu sme zistili, že vysoké trhové ceny Bitcoinu na prelome rokov 2017 a 2018 boli netradičnou výchyľkou z bežných hodnôt za sledované dvojročné obdobie. Ceny narastali do neprimeraných výšok a častokrát patrili medzi odľahlé hodnoty. To nám naznačuje, že podobný stav s extrémnym nárastom cien nie je bežnou situáciou vo vývoji kryptomeny. Ďalej sme napríklad z korelačnej matice zistili, že trhovú cenu a objem uskutočnených transakcií nad Bitcoinom priamoúmerne súvisia, a teda má zmysel sledovať ich vzájomný vývoj, ak chceme predvídať krátkodobý stav na kryptotrhu. Podobným spôsobom môžeme vytvárať mnoho ďalších vizualizácií, analyzovať ich a nachádzať v nich zaujímavé paralely a spojenia medzi rôznymi atribútmi dát.

Značnú časť diplomovej práce sme venovali vizualizácii viacrozmerných dát. Dielo Inselberga [2] o metóde paralelných súradníc nám vnuklo nápad na vytvorenie vlastnej vizualizácie s využitím piatich rozmerov z dát o kryptomene Bitcoin. Na aplikáciu sme si vybrali softvér Tableau, keďže paralelná súradnicová sústava sa nenachádza v základnom balíku funkcionalít, a zároveň Tableau ponúka priaznivé prostredie na dátovú analýzu. Vytvorili sme podrobný postup, na základe ktorého si môže študent vytvoriť vlastný graf a pomocou neho analyzovať vybrané dáta. V prílohe sme uviedli aj obrázkový postup z celej tvorby. Zaujímavosťou je, že ak by sme priložené obrázky uviedli do pohybu v rýchlom slede za sebou, tak by vznikol film.

Z paralelnej súradnicovej sústavy sme vybrali momenty, ktoré sme vyhodnotili ako zaujímavé na analyzovanie. Opäť sme si potvrdili, že objem transakcií nad Bitcoinom a jeho trhovú cenu spolu úzko súvisia. V spojení trhovej ceny s dominanciou sme nedokázali pre zvolený rozsah dátového súboru určiť jednoznačný trend. Avšak viaceré náznaky hovorili o negatívnej korelácii. To znamená, že dominancia zvyčajne klesala so stúpajúcou cenou. Skrytý význam by sme mohli nájsť v tom, že so zvyšovaním ceny začali investori odpredávať svoje podiely a investovať do iných kryptomien.

V súčasnosti sa trh s kryptomenou Bitcoin dostáva do podobného stavu, v akom sa nachádzal v prvej polovici roku 2017, pred výrazným kolísaním trhových cien. Či sa cenová situácia v blízkej budúcnosti zopakuje sa nedá jednoznačne povedať. Na podobné predpovede je nevyhnutné poznať pozadie ďalších faktorov, ktoré ovplyvňujú vývoj kryptomeny Bitcoin. Do budúcnosti navrhujeme vytvoriť nové vizualizácie s využitím poznatkov získanými z tejto práce. Rozšírením tejto témy a skombinovaním viacerých grafov budeme schopní spojiť niekoľko informácií dohromady a vytvoriť kvalitnejšie analýzy. Ako spomenul Vaganyan v svojej knihe [1], pre manažéra je dôležité vidieť kombinované vizualizácie informácií, aby bol schopný optimalizovať svoje rozhodnutia.

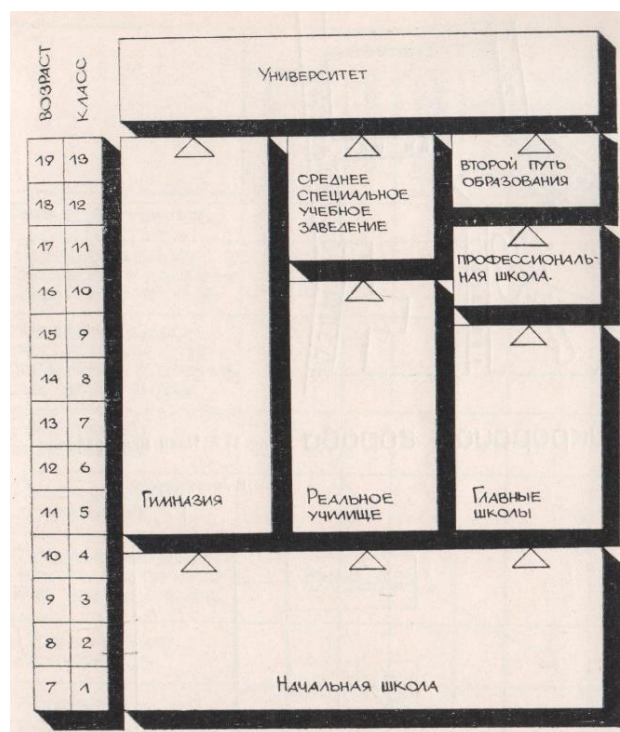
Zoznam použitej literatúry

- [1] G. A. VAGANYAN, *Mašinnaja grafika v upravlenii*, Jerevan: Ajastan, 1985, s. 145-169.
- [2] A. INSELBERG, *Parallel coordinates - visual multidimensional geometry and its applications*, New York: Springer, 2009, s. 1-75.
- [3] „CoinMarketCap,“ CoinMarketCap, 2019. [Online]. Dostupné na: www.coinmarketcap.com. [Cit. 14. január 2019].
- [4] „Bitcoins in circulation,“ Blockchain, 2017. [Online]. Dostupné na: www.blockchain.com/charts/total-bitcoins. [Cit. 14. január 2019].
- [5] „Tableau for students,“ Tableau, 2019. [Online]. Dostupné na: www.tableau.com/academic/students. [Cit. 1. marec 2019].
- [6] „What is R?,“ The R Foundation, [Online]. Dostupné na: www.r-project.org/about.html. [Cit. 1. marec 2019].
- [7] B. H. MCCORMICK, T. A. DEFANTI a M. D. BROWN, *Visualization in scientific computing*, New York: ACM SIGGRAPH, 1987, s. 3-7.
- [8] E. RUŽICKÝ a A. FERKO, *Počítačová grafika a spracovanie obrazu*, Bratislava: SAMOSATO, 2012, s. 1-3.
- [9] G. DOMIK, „Tutorial on visualization,“ 29. marec 1999. [Online]. Dostupné na: <http://www6.uniovi.es/hypvis/domik/folien.html>. [Cit. 14. január 2019].
- [10] E. R. TUFTE, *The visual display of quantitative information*, Cheshire: Graphics Press, 2001, s. 10-41.
- [11] E. H. CHI, „A taxonomy of visualization techniques using the data state reference model,“ 2000. [Online]. Dostupné na: <https://www.ics.uci.edu/~kobsa/courses/ICS280/InfoViz2000/ed-chi.pdf>. [Cit. 12. február 2019].

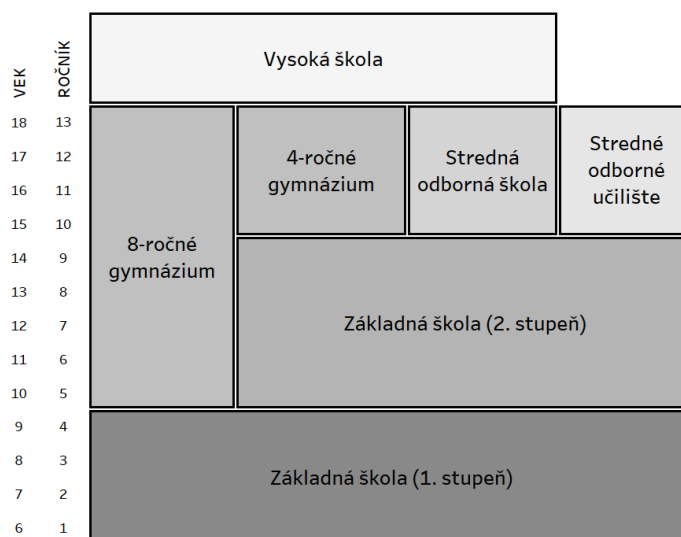
- [12] R. LENGLER a M. J. EPPLER, „*Towards a periodic table of visualization methods for management*,“ 2007. [Online]. Dostupné na: www.visual-literacy.org/periodic_table/periodic_table.pdf. [Cit. 28. február 2019].
- [13] R. LENGLER a M. J. EPPLER, „*A periodic table of visualization methods*,“ Visual Literacy, [Online]. Dostupné na: www.visual-literacy.org/periodic_table/periodic_table.html. [Cit. 28. február 2019].
- [14] D. ROAM, *The back of the napkin: solving problems and selling ideas with pictures*, New York: Portfolio, 2008, s. 3-144.
- [15] F. LAMOŠ a R. POTOCKÝ, *Pravdepodobnosť a matematická štatistika*, Bratislava: Univerzita Komenského Bratislava, 1998, s. 11-136.
- [16] H. C. THODE, *Testing for normality*, New York: Marcel Dekker, 2002, s. 15-251.
- [17] S. S. SHAPIRO a M. B. WILK, „*An analysis of variance test for normality (complete samples)*,“ *Biometrika*, zv. 52, vyd. 3/4, s. 591-611, 1965.
- [18] G. SHURKHOVETSKYY, N. ANDRIENKO, G. ANDRIENKO a G. FUCHS, „*Data abstraction for visualizing large time series*,“ *Computer Graphics Forum*, s. 1-20, 2017.
- [19] „*UNIC Blockchain Initiative*,“ University of Nicosia, 2017. [Online]. Dostupné na: www.digitalcurrency.unic.ac.cy. [Cit. 4. marec 2019].
- [20] G. VAN DE VEN, „*Exploratory data analysis*,“ 10. september 2010. [Online]. Dostupné na: www.stat.berkeley.edu/~gido/Exploratory%20Data%20Analysis.pdf. [Cit. 6. marec 2019].
- [21] M. G. KENDALL, *Rank correlation methods*, London: Griffin, 1970, s. 1-66.
- [22] H. A. DAVID, H. O. HARTLEY a E. S. PEARSON, „*The distribution of the ratio, in a single normal sample, of range to standard deviation*,“ *Biometrika*, zv. 41, vyd. 3/4, s. 482-493, 1954.

Príloha A

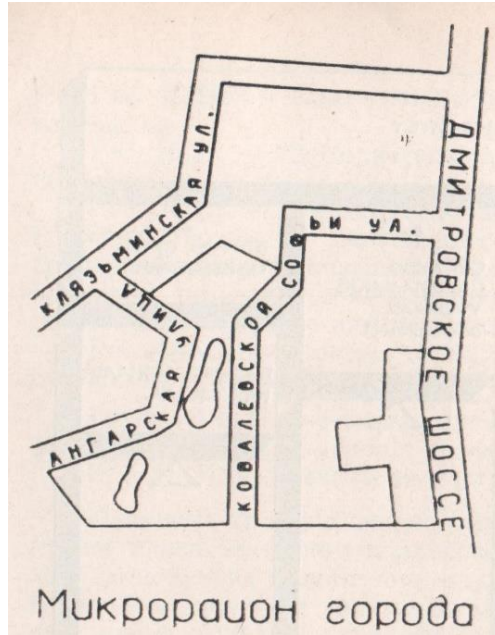
Novodobé ekvivalenty k vizualizáciám informácií



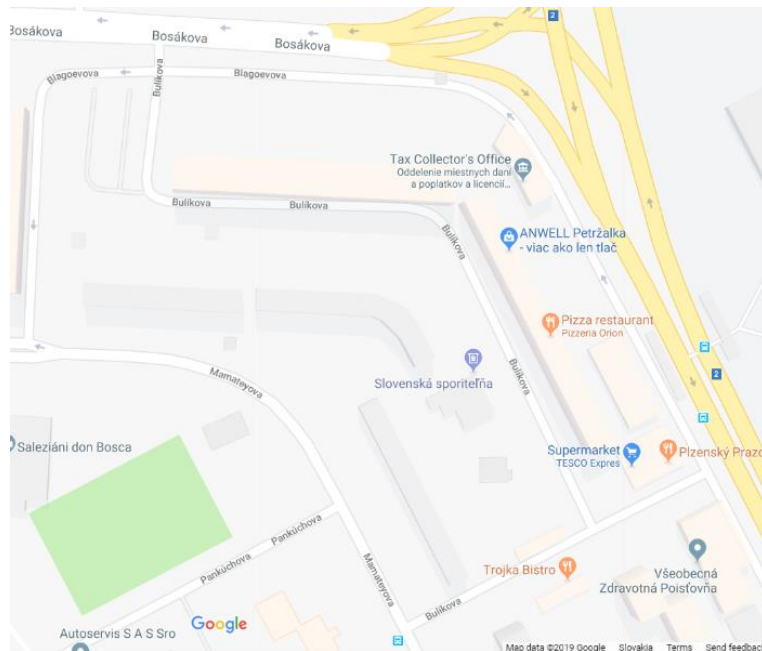
Obr. A1 - Vzdelávací systém [1]



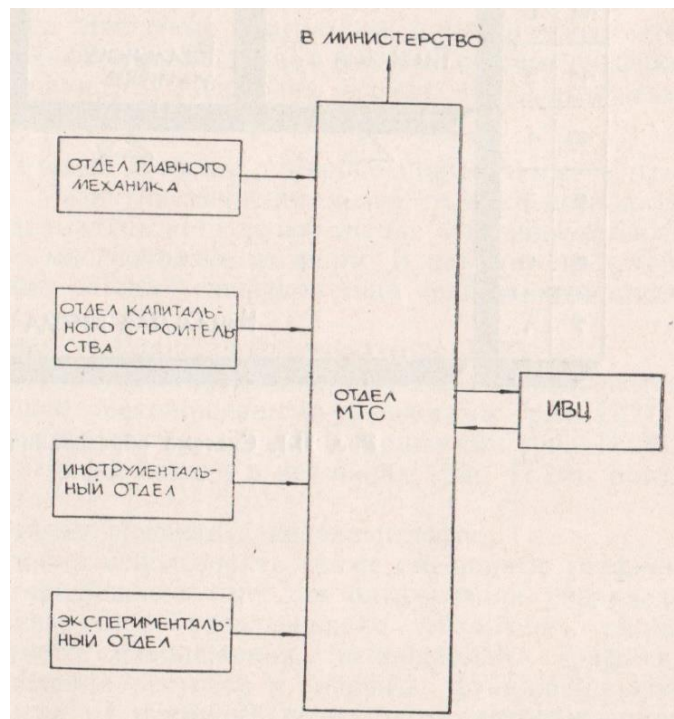
Obr. A2 - Vzdelávací systém vytvorený v Tableau



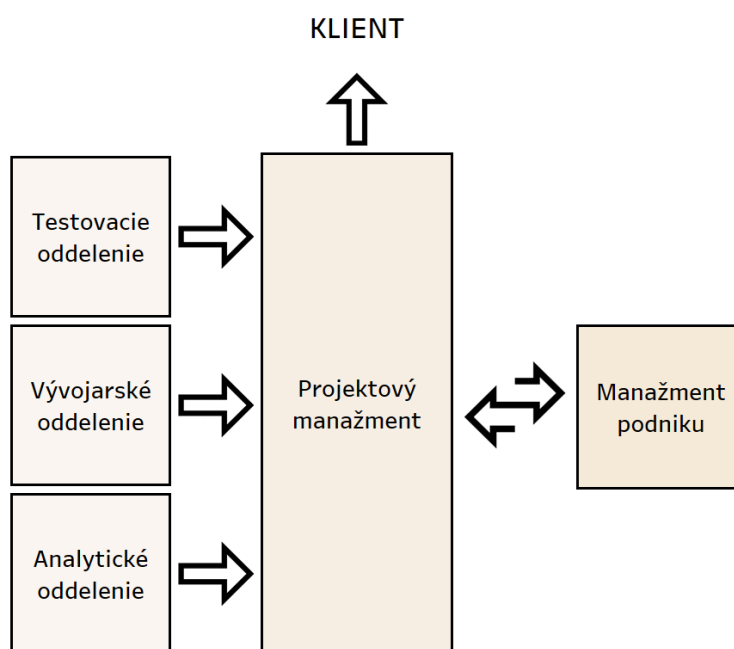
Obr. A3 - Plán sídliska [1]



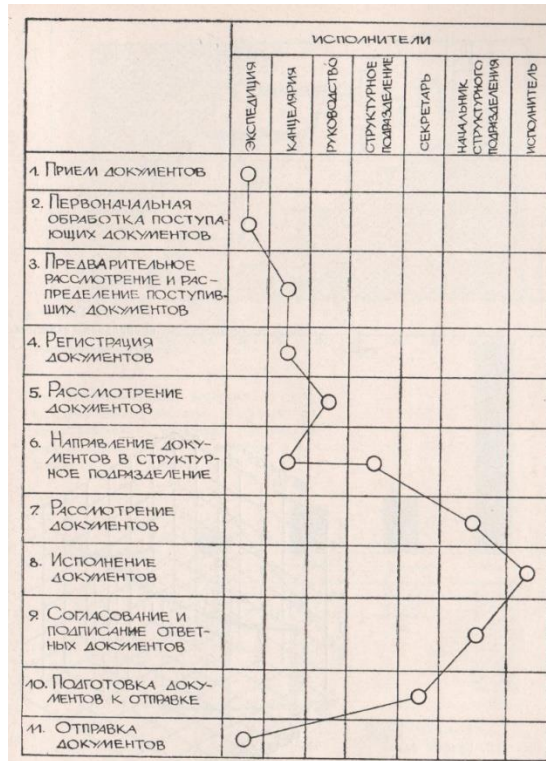
Obr. A4 - Plán sídliska z Google Maps



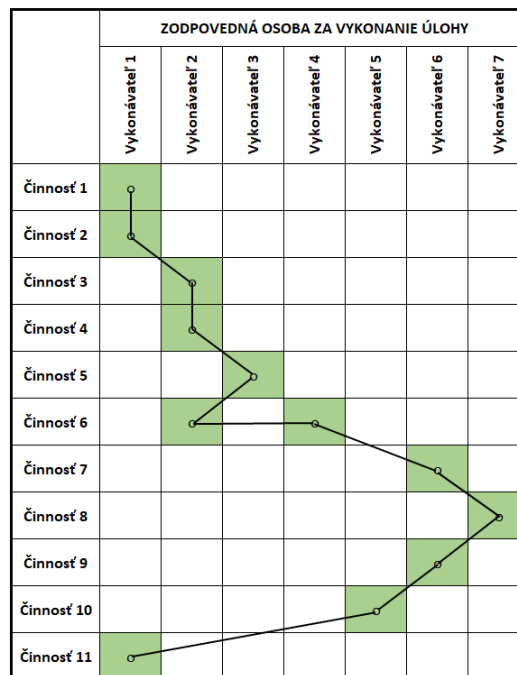
Obr. A5 - Schéma informačného toku v podniku [1]



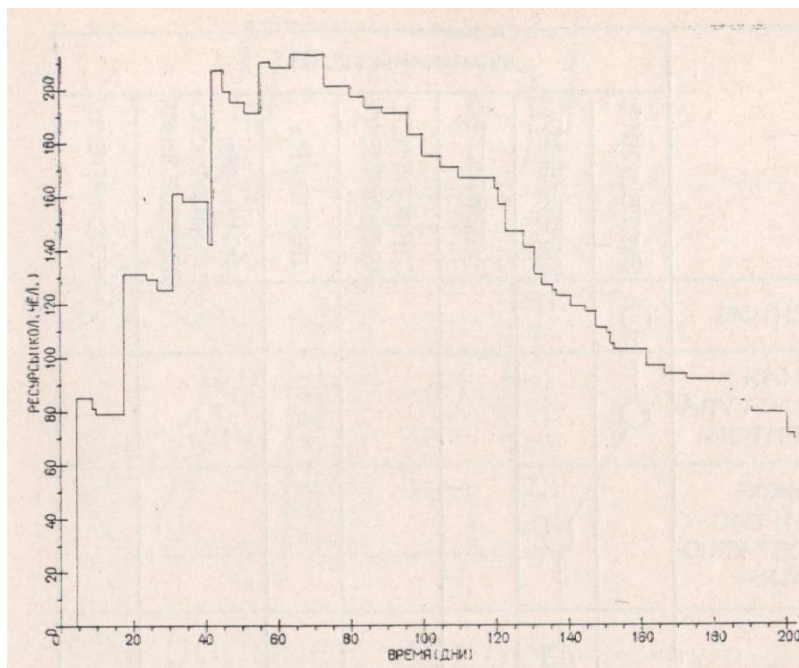
Obr. A6 - Schéma informačného toku v podniku vytvorená v Tableau



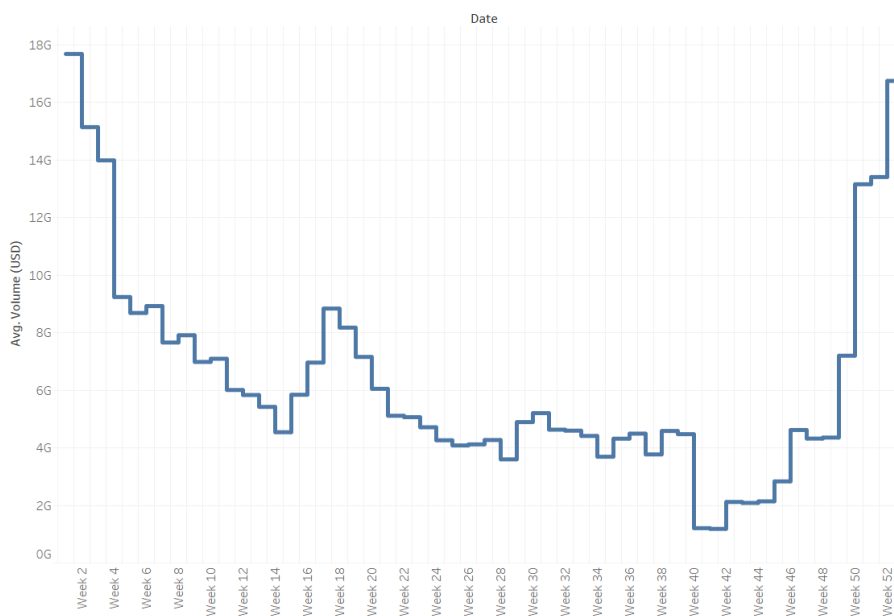
Obr. A7 - Diagram pohybu dokumentov a procesov [1]



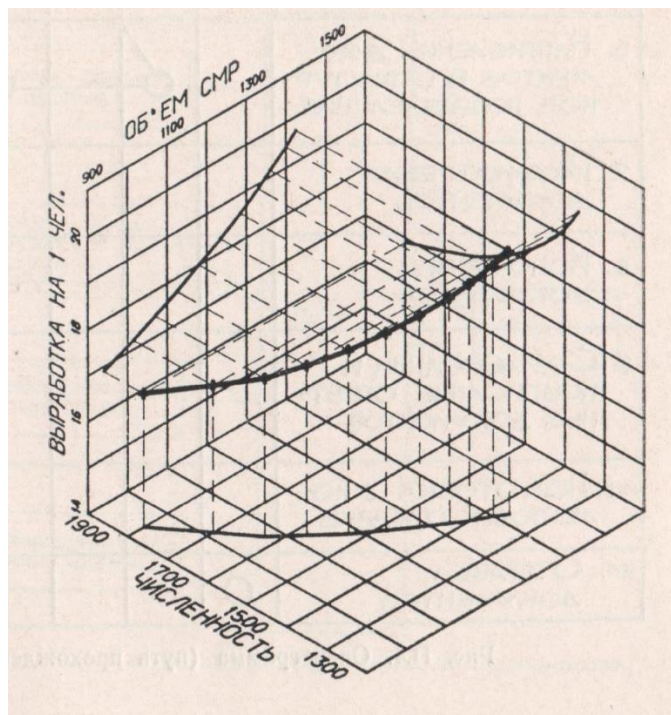
Obr. A8 - Diagram pohybu dokumentov vytvorený v MS Excel



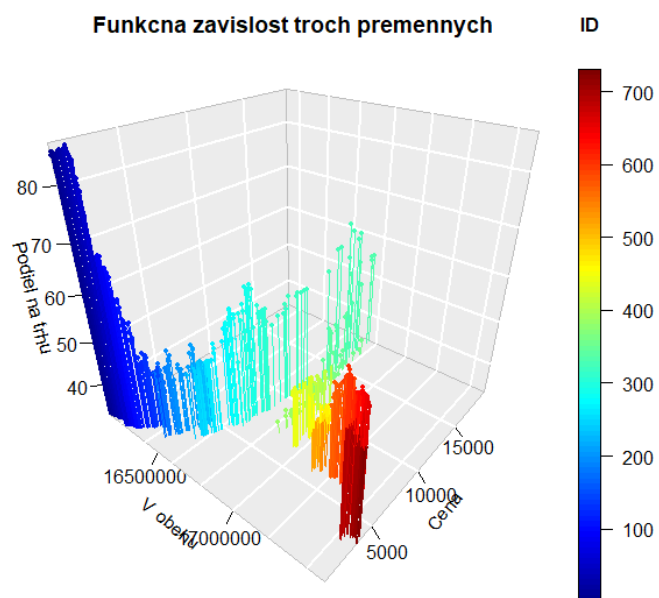
Obr. A9 - Graf funkčních závislostí dvou proměných [1]



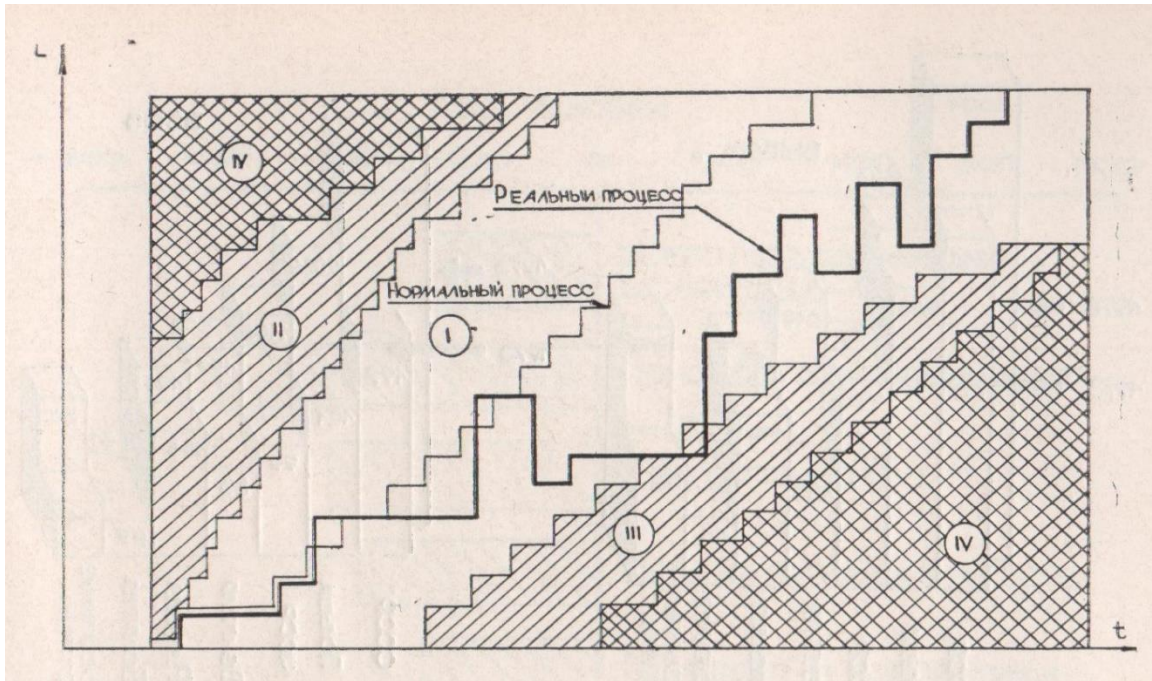
Obr. A10 - Graf funkčních závislostí dvou proměnných vytvořený v Tableau



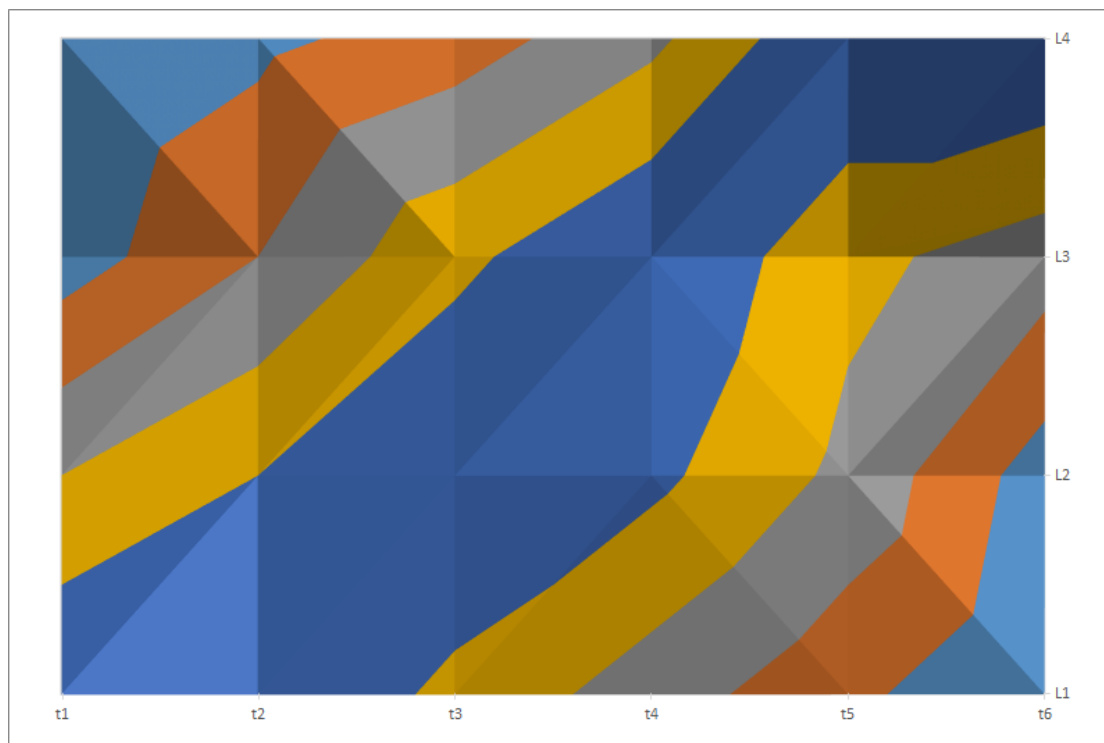
Obr. A11 - Graf funkčných závislostí troch premenných [1]



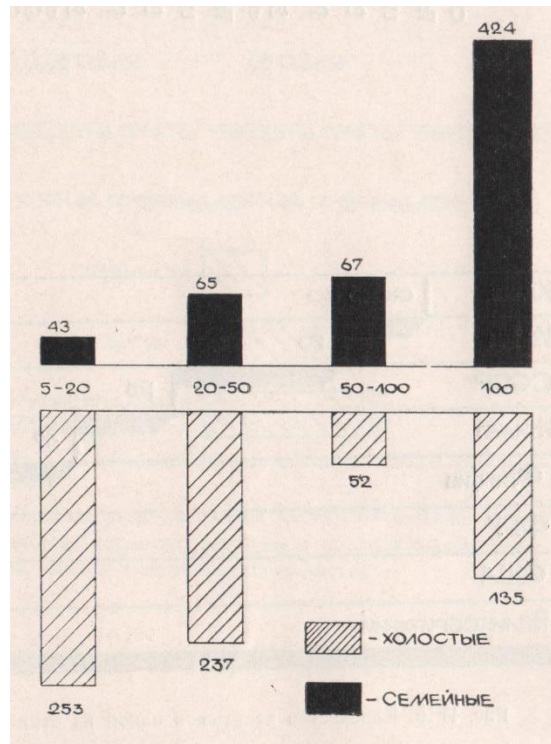
Obr. A12 - Graf funkčných závislostí troch premenných vytvorený v R



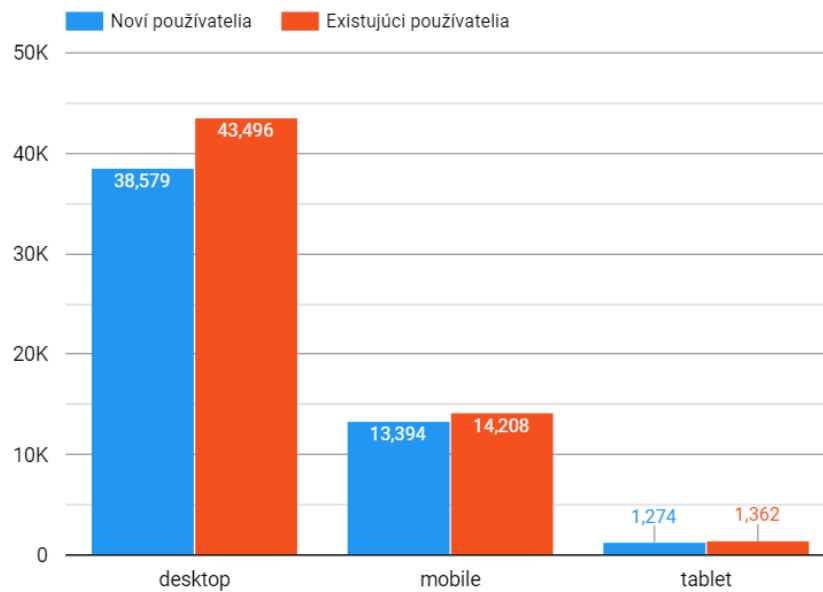
Obr. A13 - Integrovaný model riadenia informácií



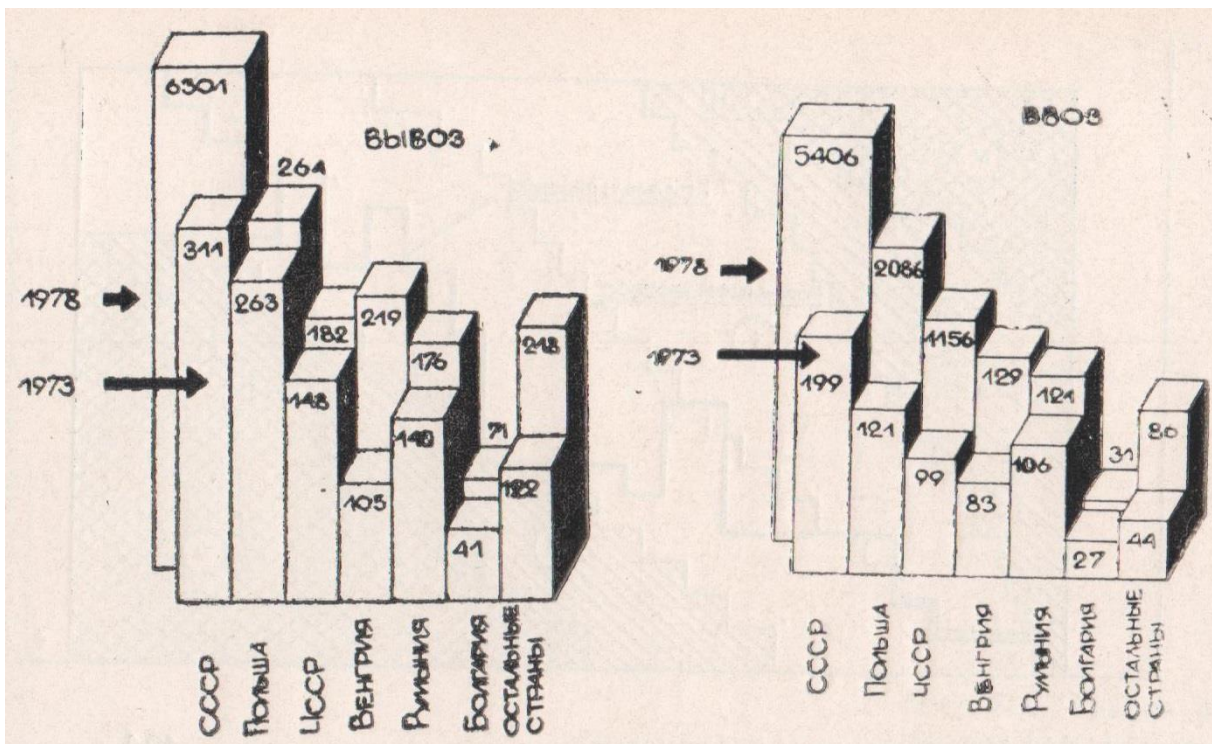
Obr. A14 - Povrchový graf vytvorený v MS Excel



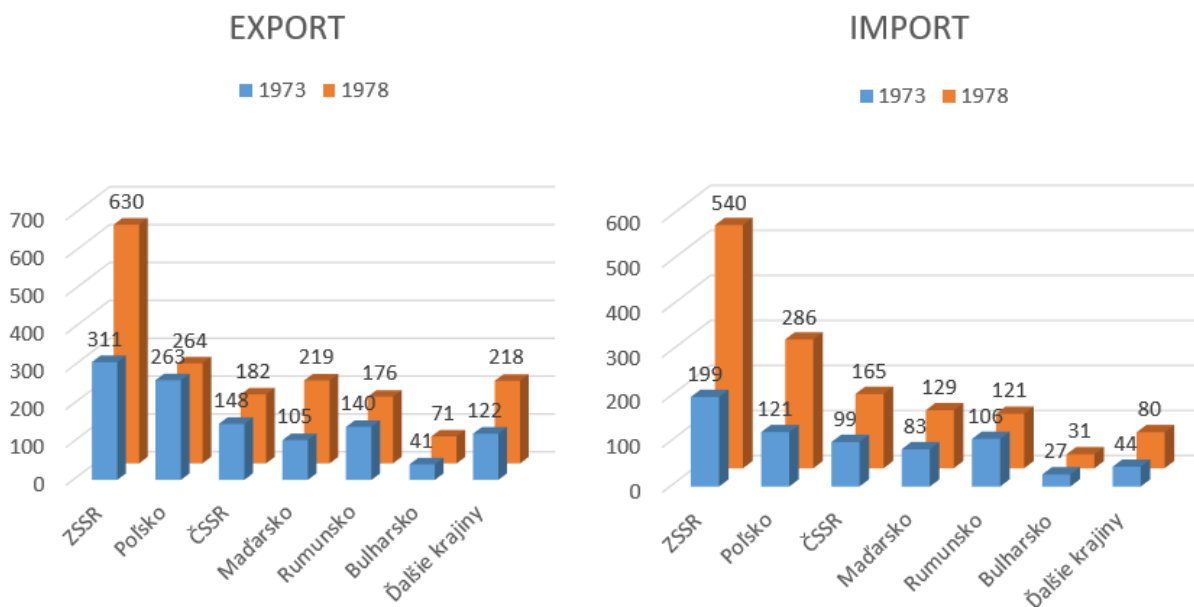
Obr. A15 - Stĺpcový graf [1]



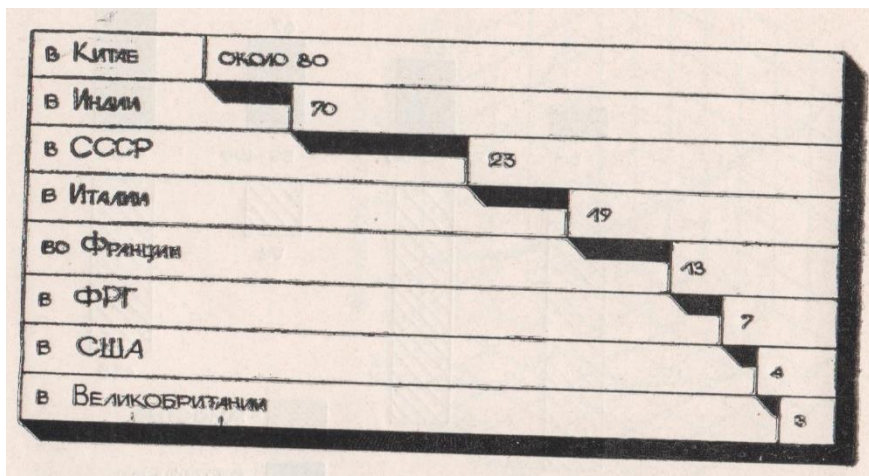
Obr. A16 - Stĺpcový graf vytvorený v Google Data Studio



Obr. A17 - Priestorový stĺpcový graf [1]



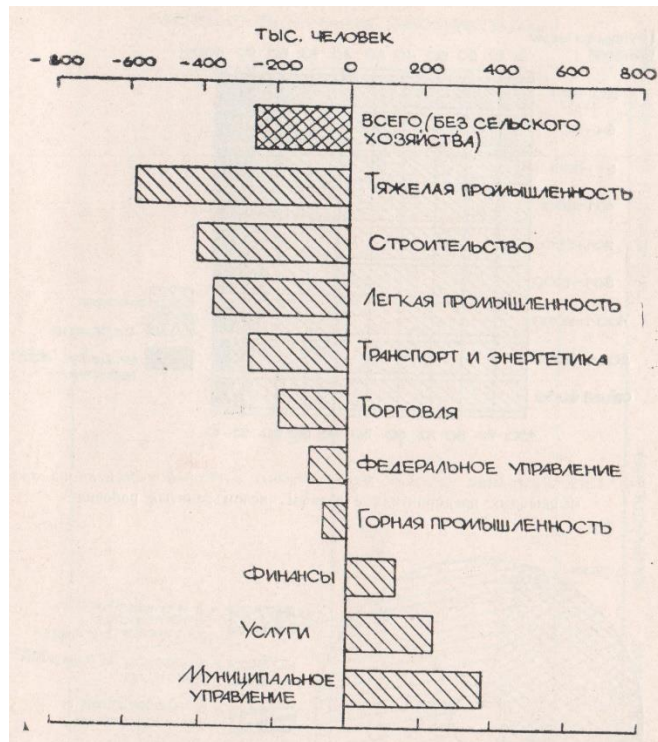
Obr. A18 - Priestorový stĺpcový graf vytvorený v MS Excel



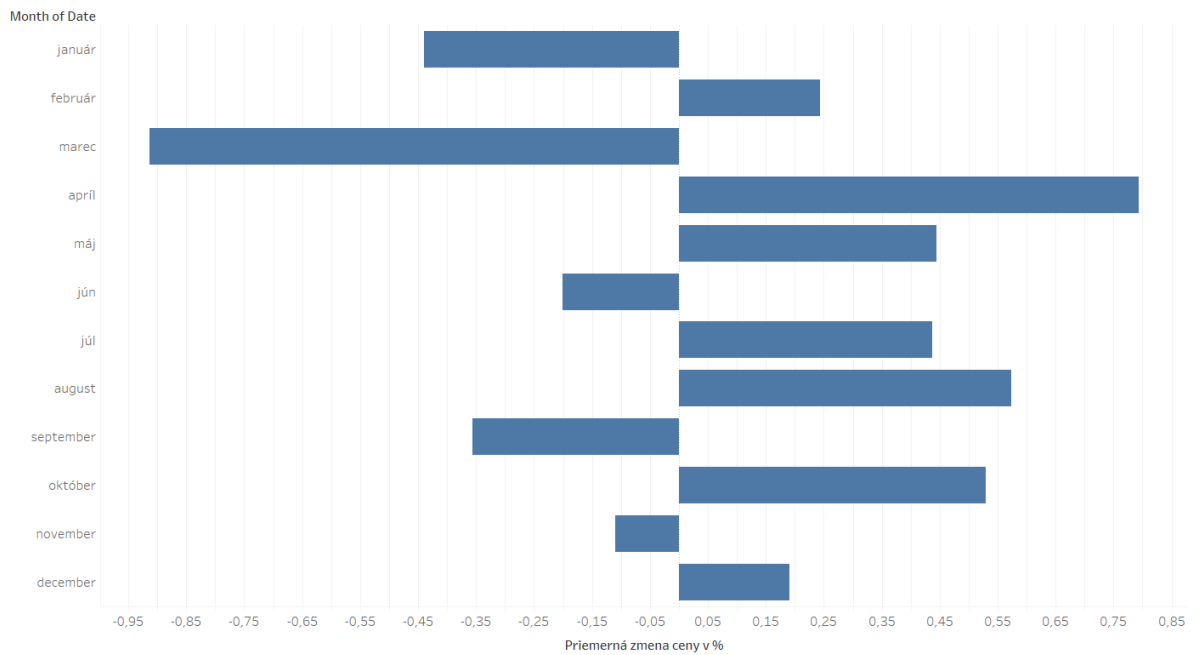
Obr. A19 - Pruhový graf [1]



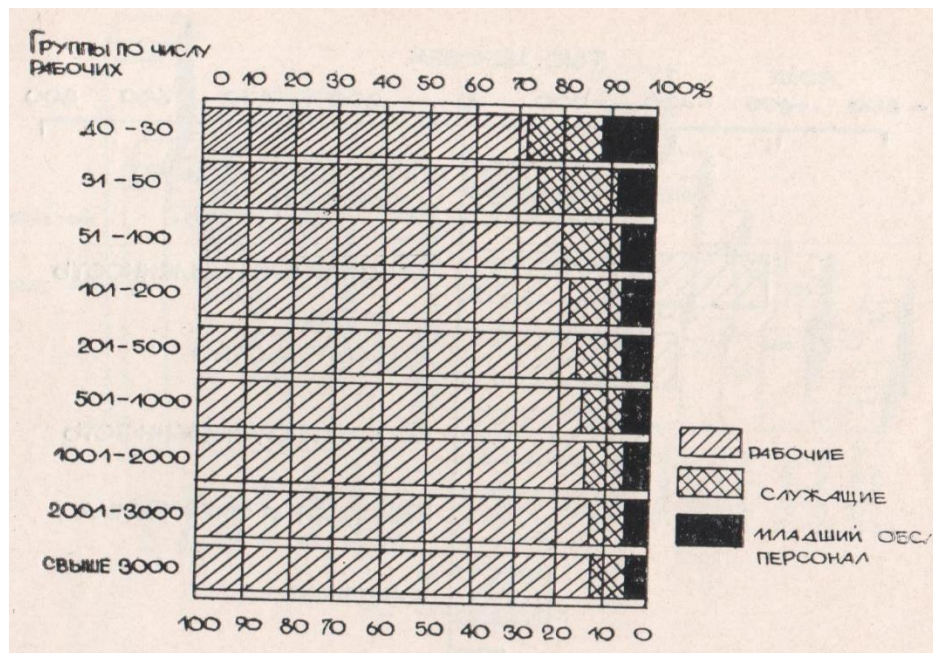
Obr. A20 - Pruhový graf vytvorený v MS Excel



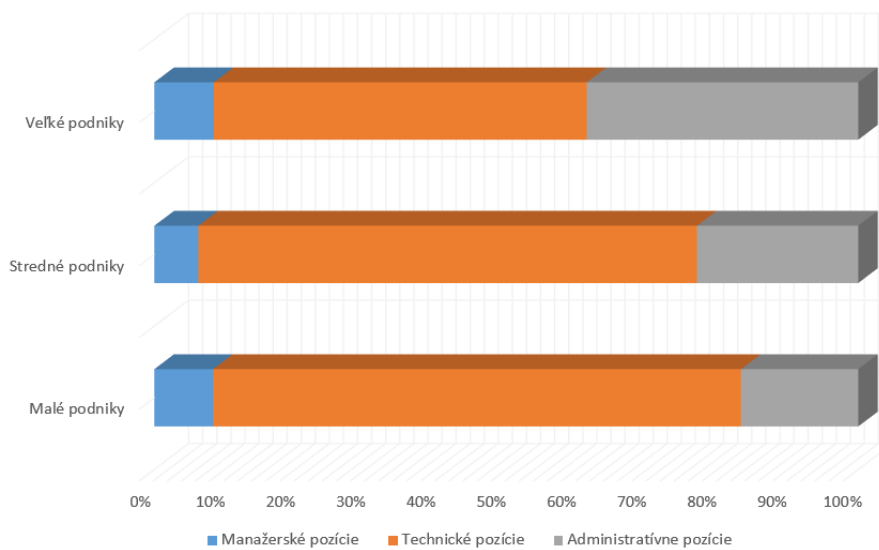
Obr. A21 - Obojstranný pruhový graf [1]



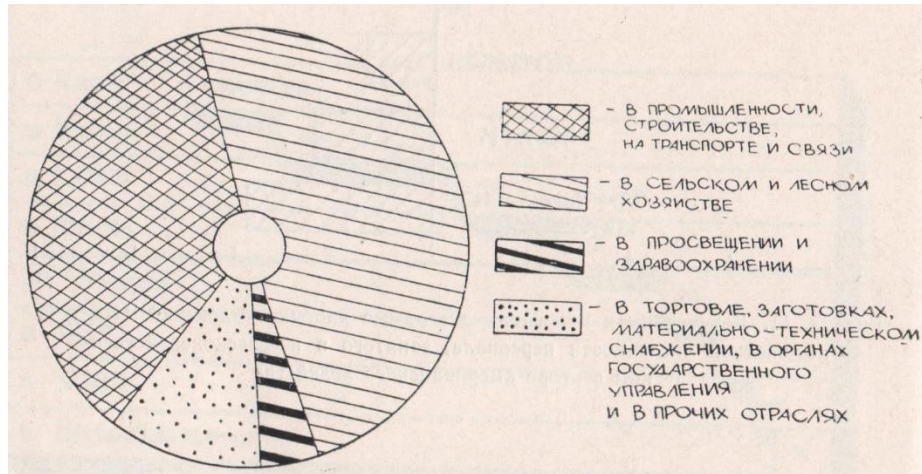
Obr. A22 - Obojstranný pruhový graf vytvorený v Tableau



Obr. A23 - Stopercentný skladaný pruhový graf [1]

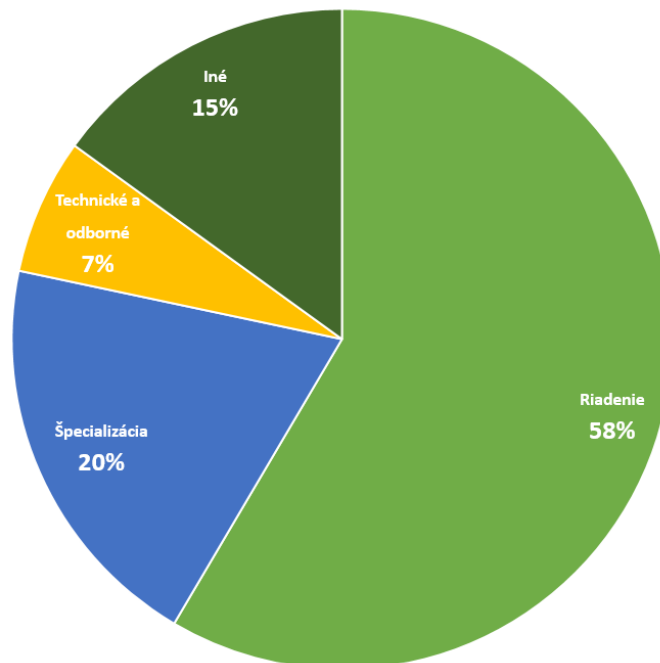


Obr. A24 - Stopercentný skladaný pruhový graf vytvorený v MS Excel

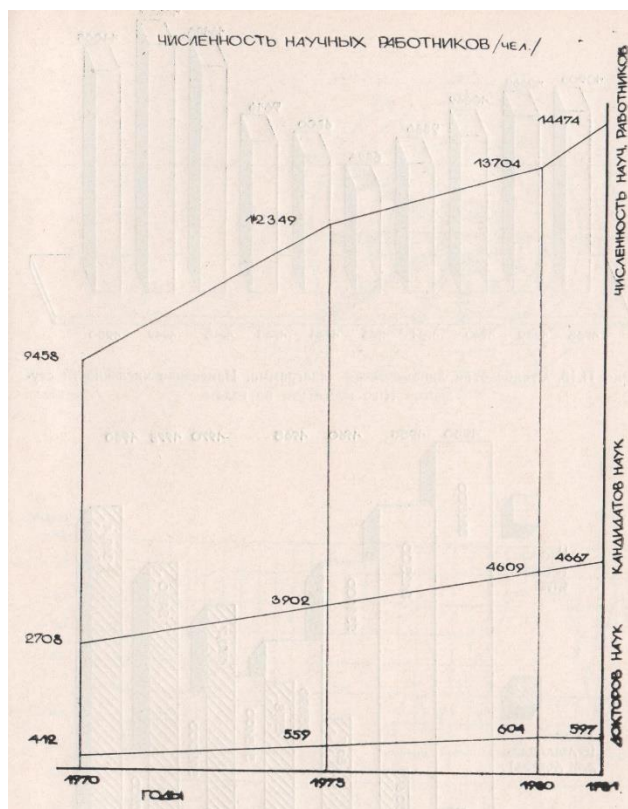


Obr. A25 - Koláčový graf [1]

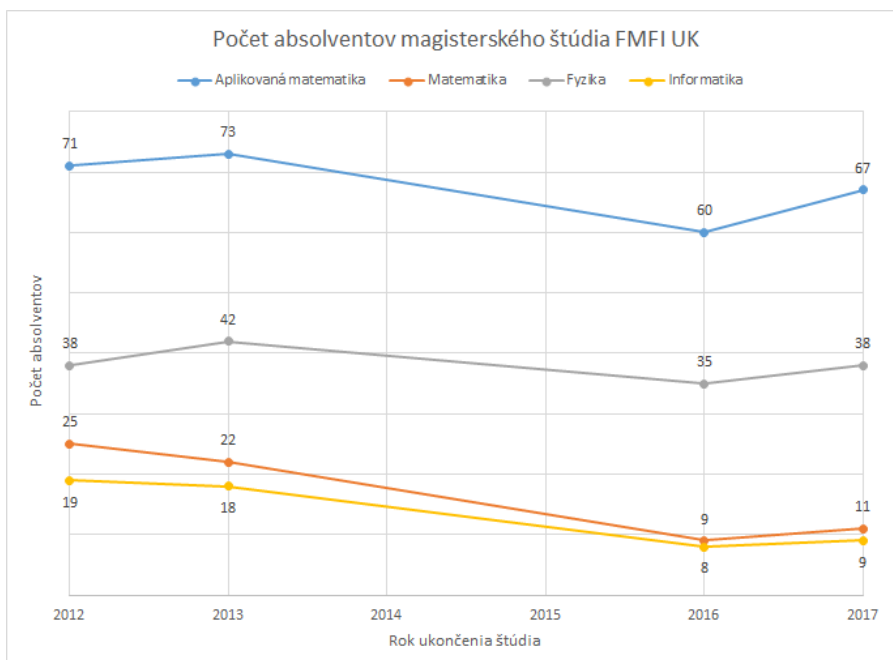
NAJLEPŠIE PLATENÉ ZAMESTNANIA



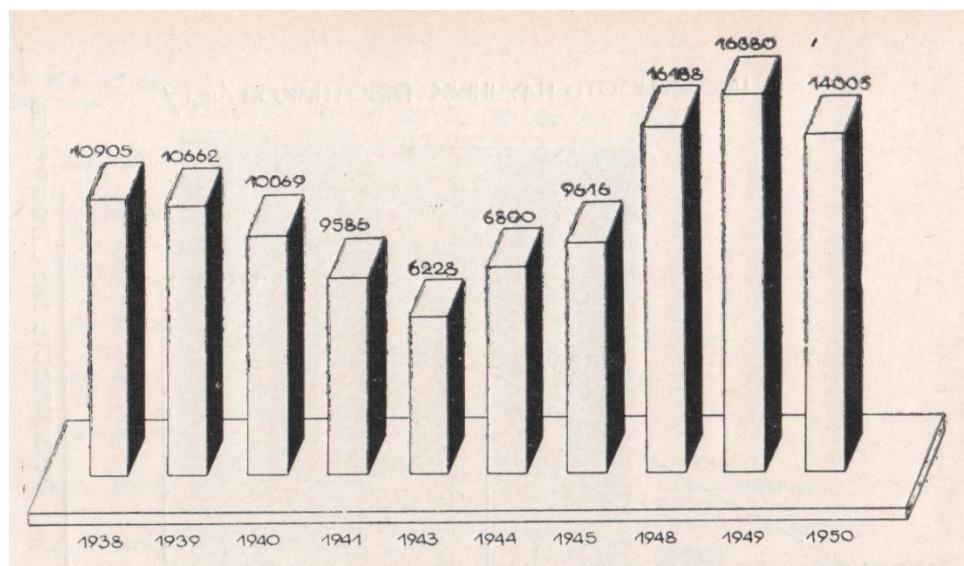
Obr. A26 - Koláčový graf vytvorený v MS Excel



Obr. A27 - Lineárny súradnicový graf [1]

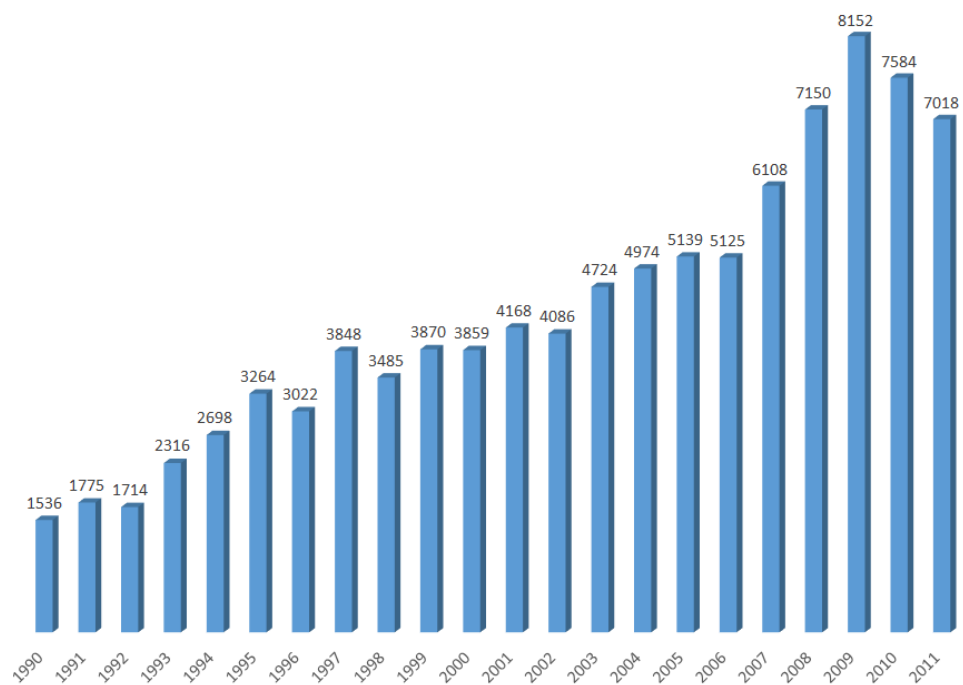


Obr. A28 - Lineárny súradnicový graf vytvorený v MS Excel

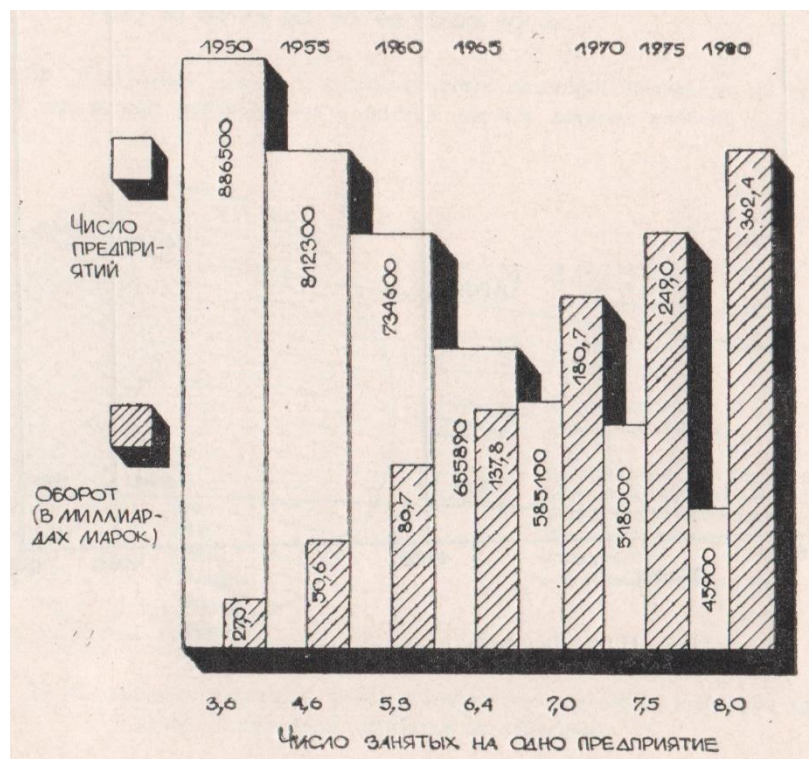


Obr. A29 - Stĺpcový dynamický graf zmeny jedného parametra [1]

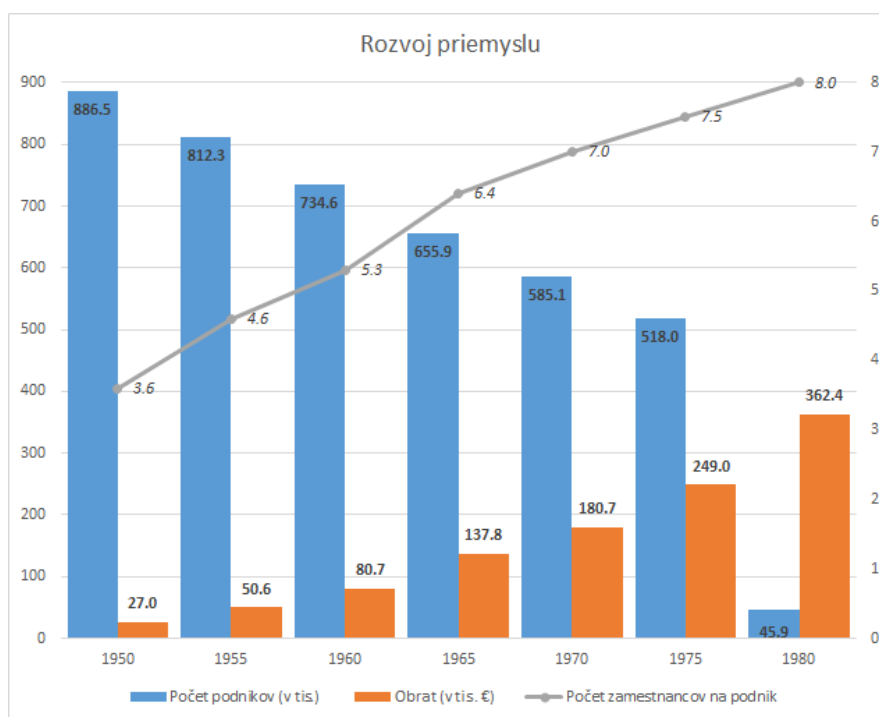
Absolventi bakalárskeho, magisterského a doktorského štúdia na UK



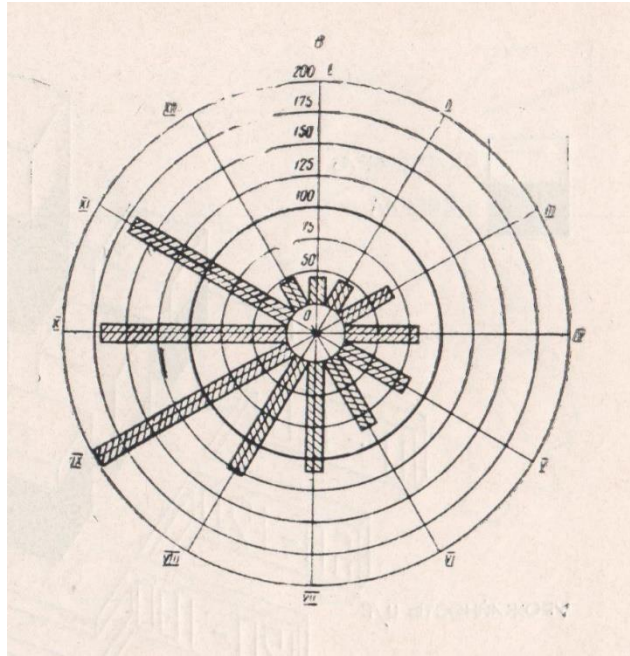
Obr. A30 - Stĺpcový dynamický graf zmeny jedného parametra vytvorený v MS Excel



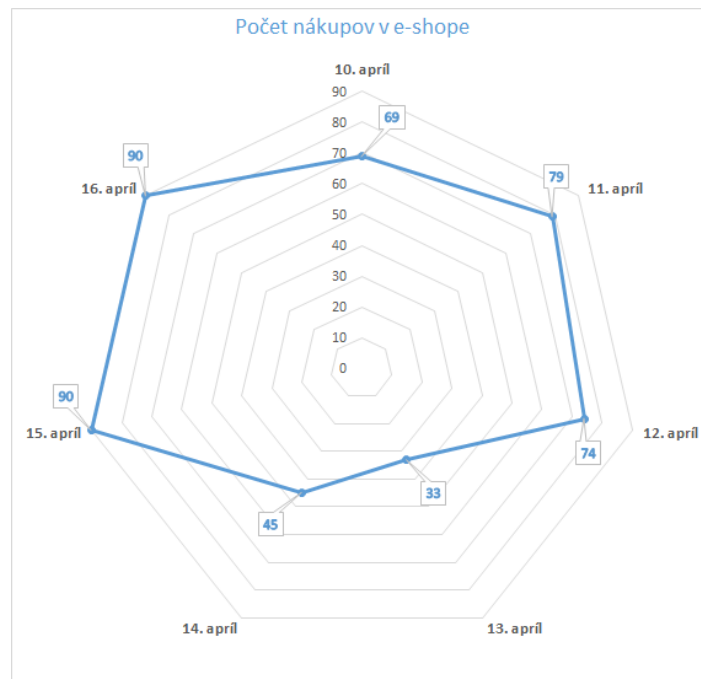
Obr. A31 - Stĺpcový dynamický graf zmeny viacerých parametrov [1]



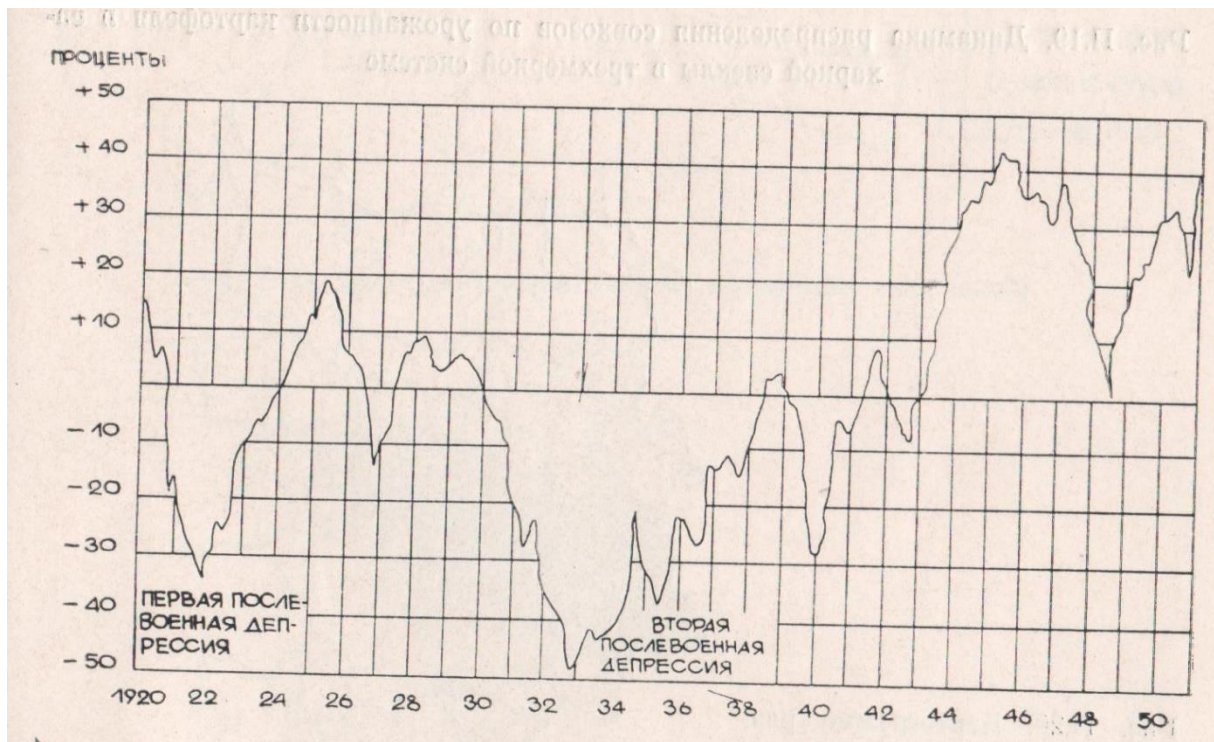
Obr. A32- Kombinovaný dynamický graf zmeny viacerých parametrov vytvorený v MS Excel



Obr. A33- Kruhový graf [1]

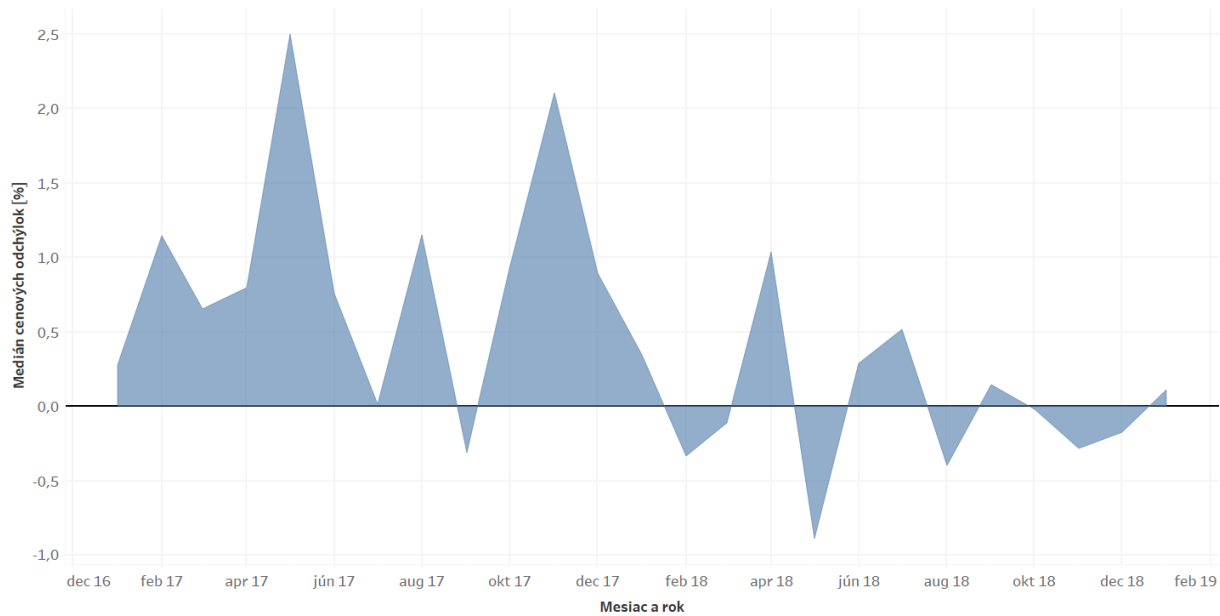


Obr. A34- Radarový graf vytvořený v MS Excel

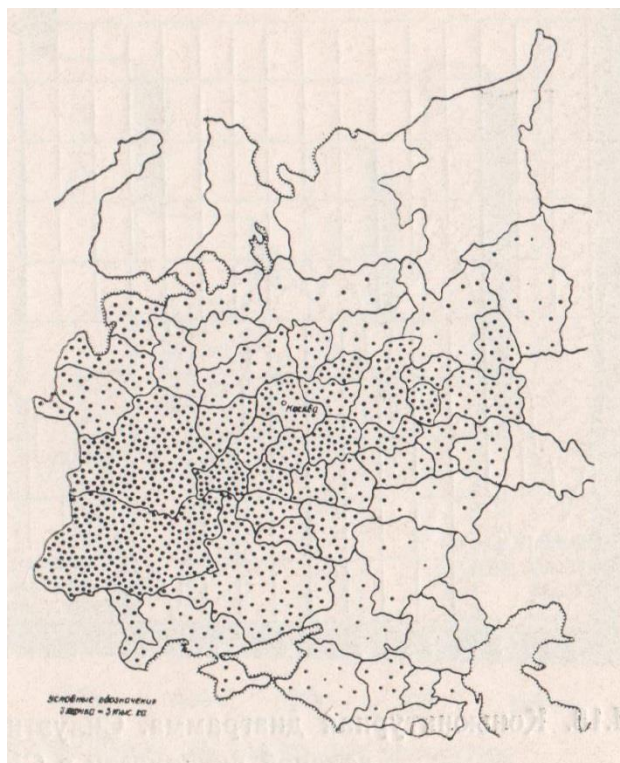


Obr. A35 - Graf trhu zobrazujúci percentuálnu zmenu parametra [1]

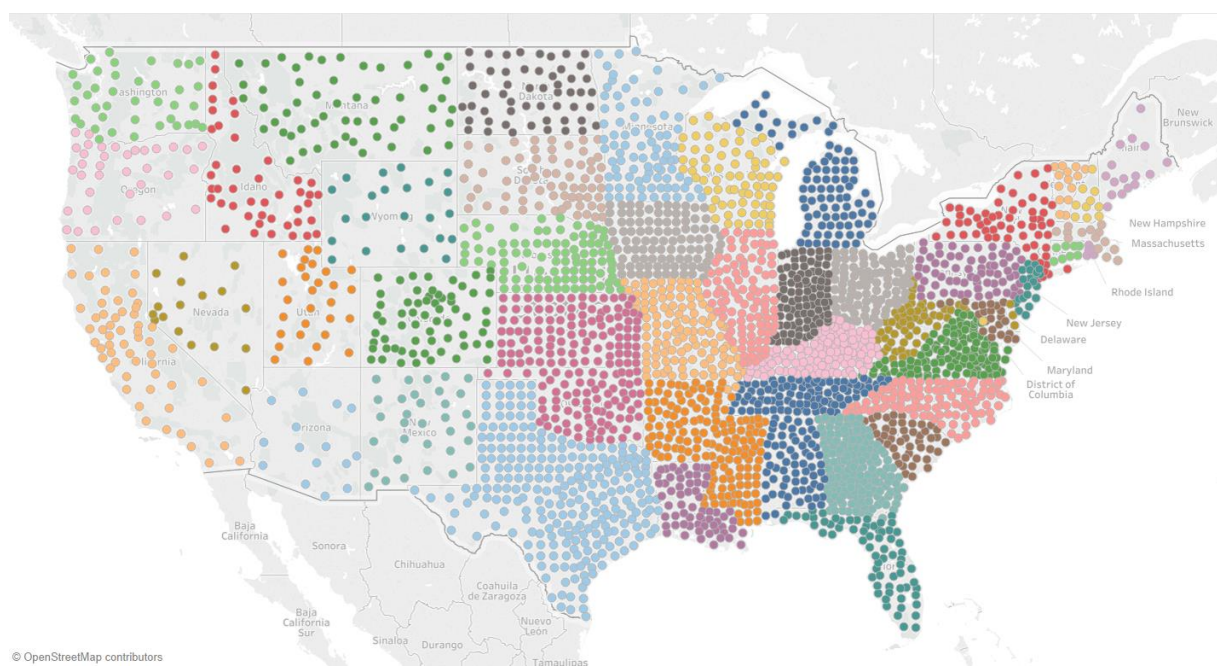
Zmena ceny Bitcoinu za jednotlivé mesiace



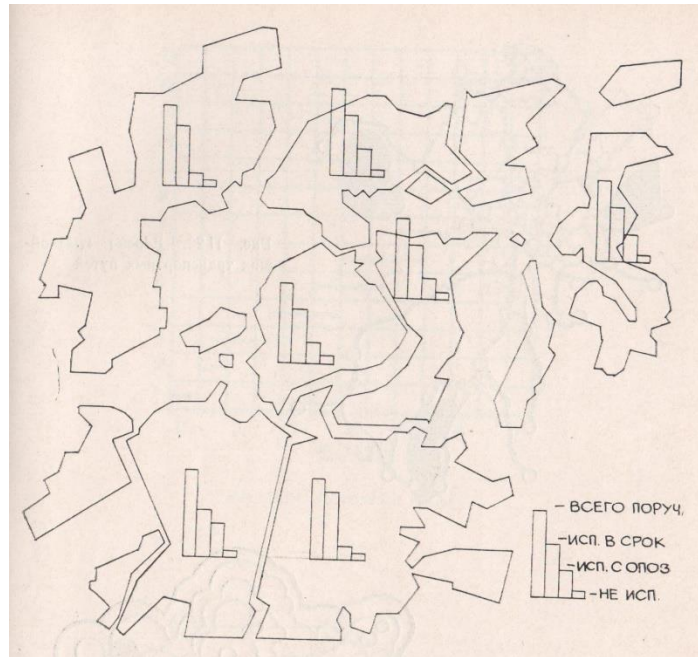
Obr. A36 - Graf zobrazujúci percentuálnu zmenu parametra vytvorený v Tableau



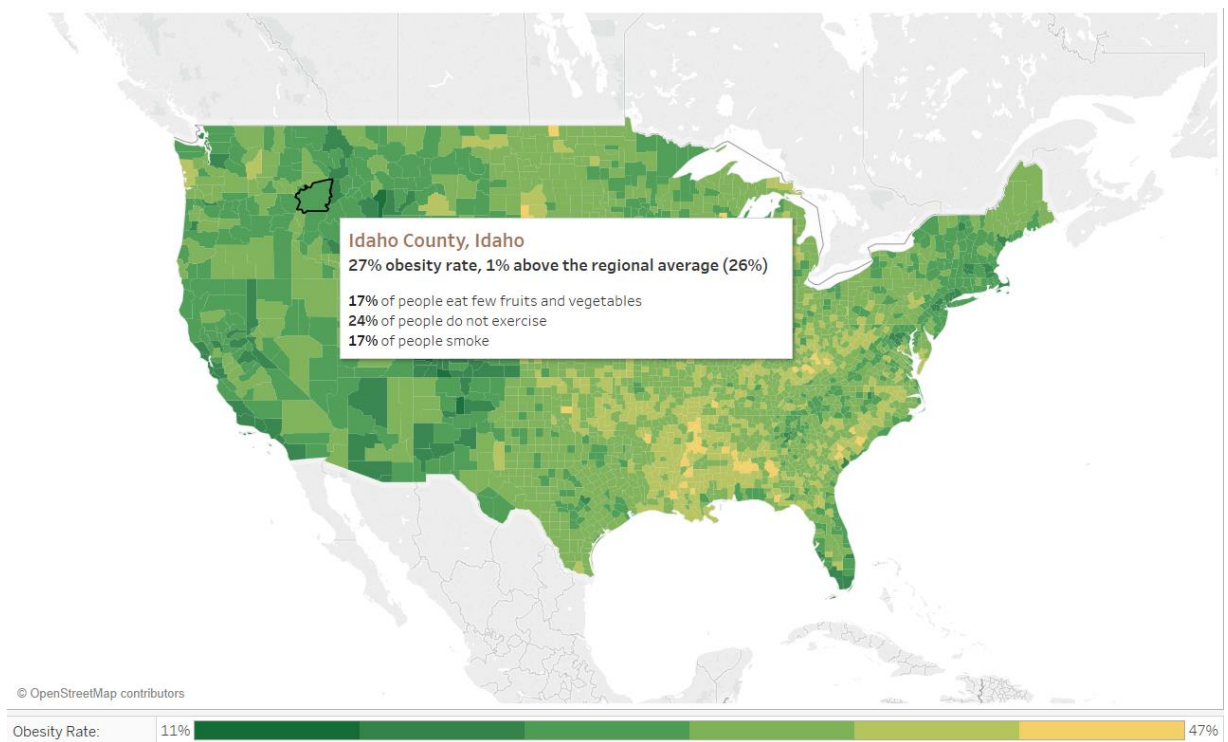
Obr. A37 - Mapa zasiatych plôch [1]



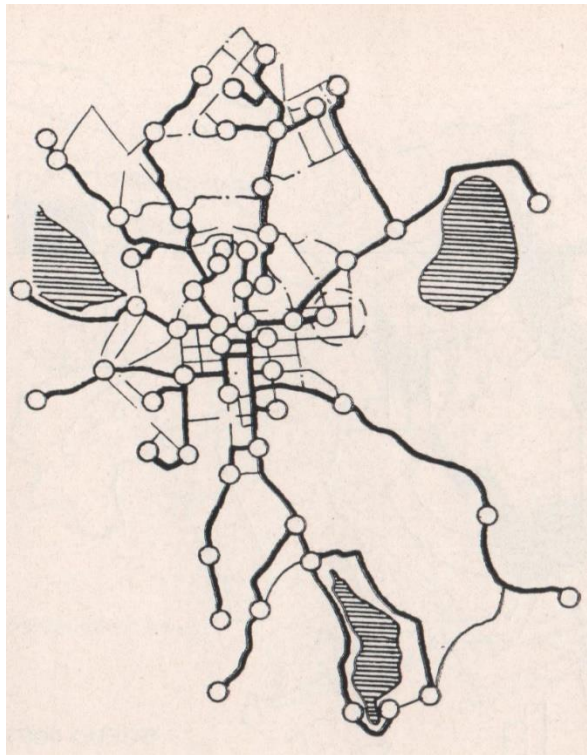
Obr. A38 - Mapa krajov v jednotlivých štátoch vytvorená v Tableau



Obr. A39 - Analyticko-štatistická mapa [1]



Obr. A40 - Analyticko-štatistická mapa vytvorená v Tableau



Obr. A41 - Mapa najkratších ciest [1]

Digital Park I, 851 01 Petržalka
 FMFI UK, pavilón matematiky, Mlynská c

Leave now OPTIONS

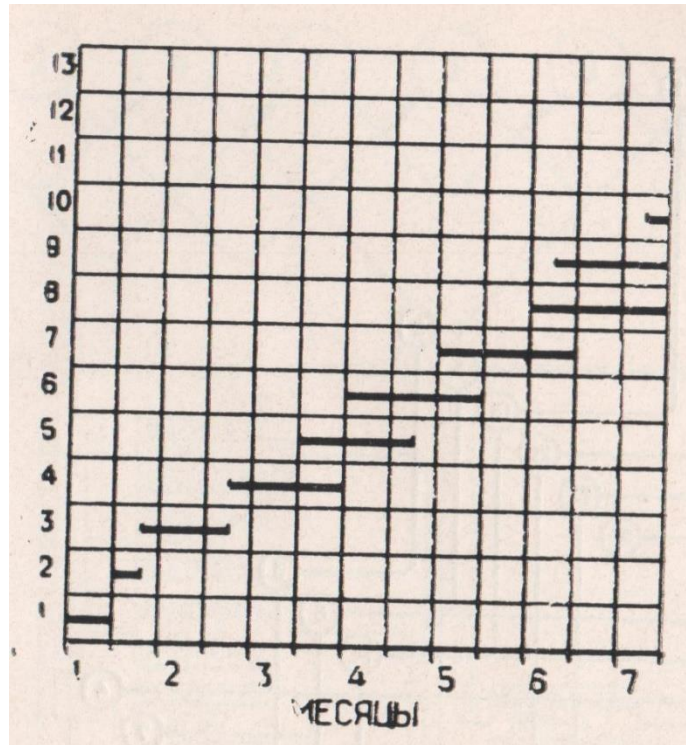
Send directions to your phone

7:40 PM–8:16 PM	36 min
7:46 PM from Dvory	
15 min every 10 min	
DETAILS	
7:46 PM–8:16 PM	30 min
7:42 PM–8:16 PM	34 min
7:40 PM–8:21 PM	41 min

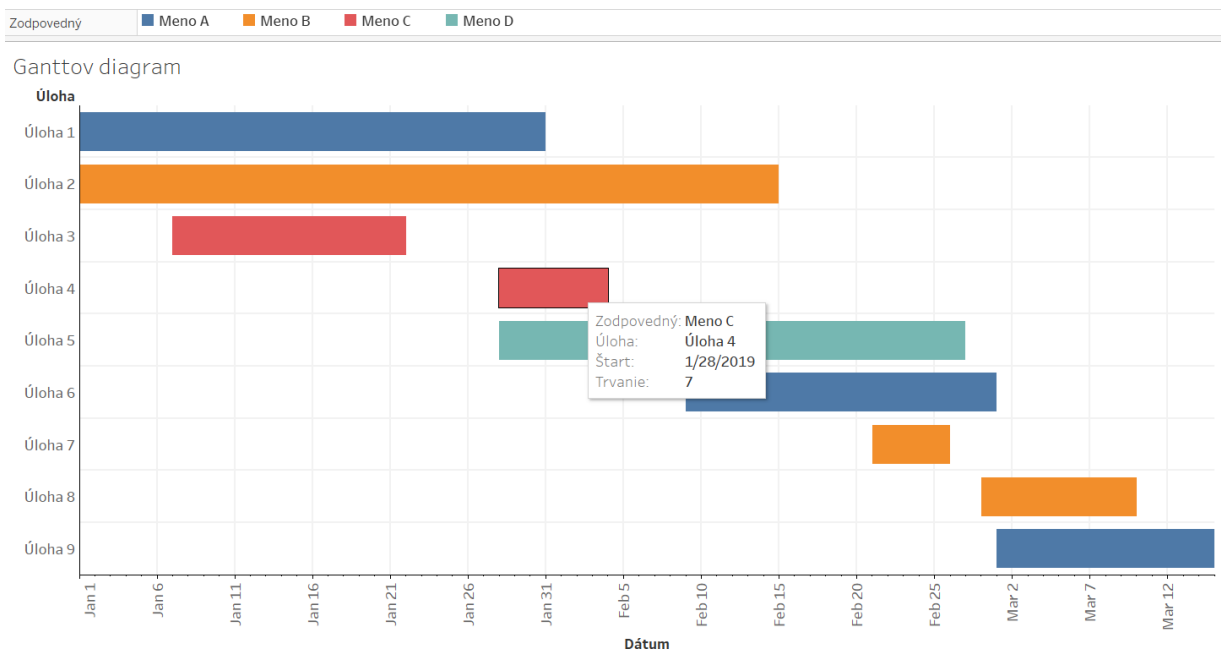
SCHEDULE EXPLORER

Map data ©2019 Google

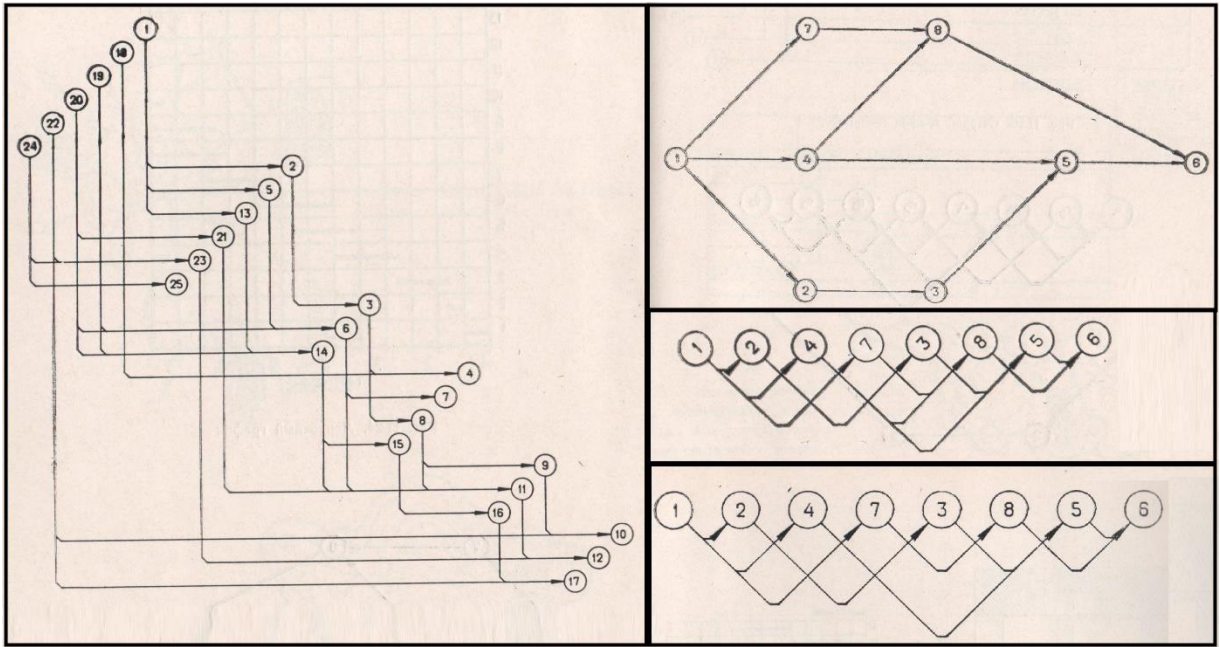
Obr. A42 - Mapa najkratších ciest z Google Maps



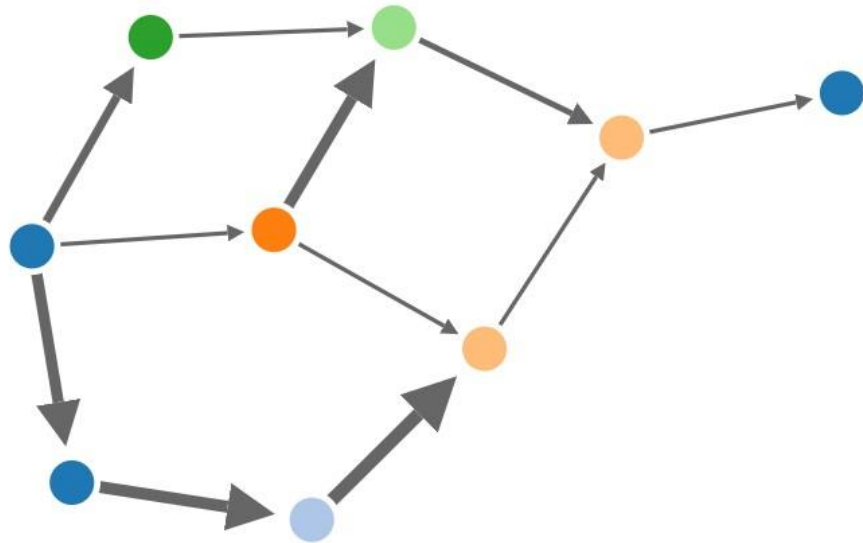
Obr. A43 - Ganttov diagram [1]



Obr. A44 - Ganttov diagram vytvorený v Tableau



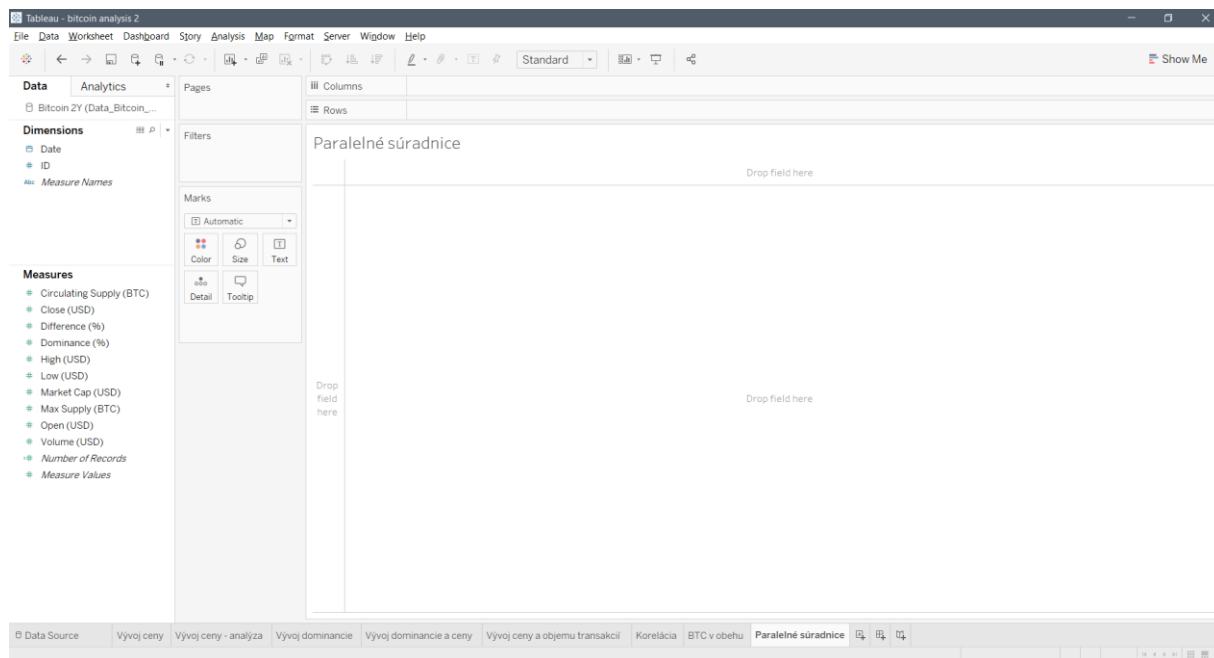
Obr. A45 - Rôzne druhy sieťových grafov [1]



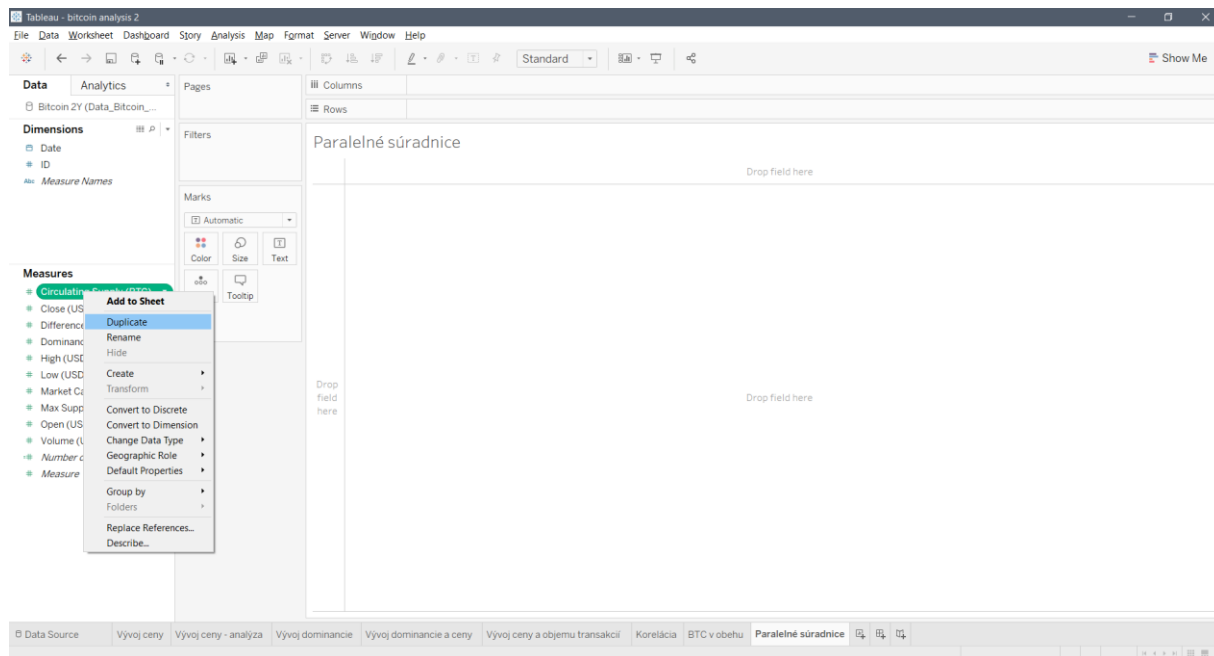
Obr. A46 - Príklad sieťového grafu vytvoreného v R

Príloha B

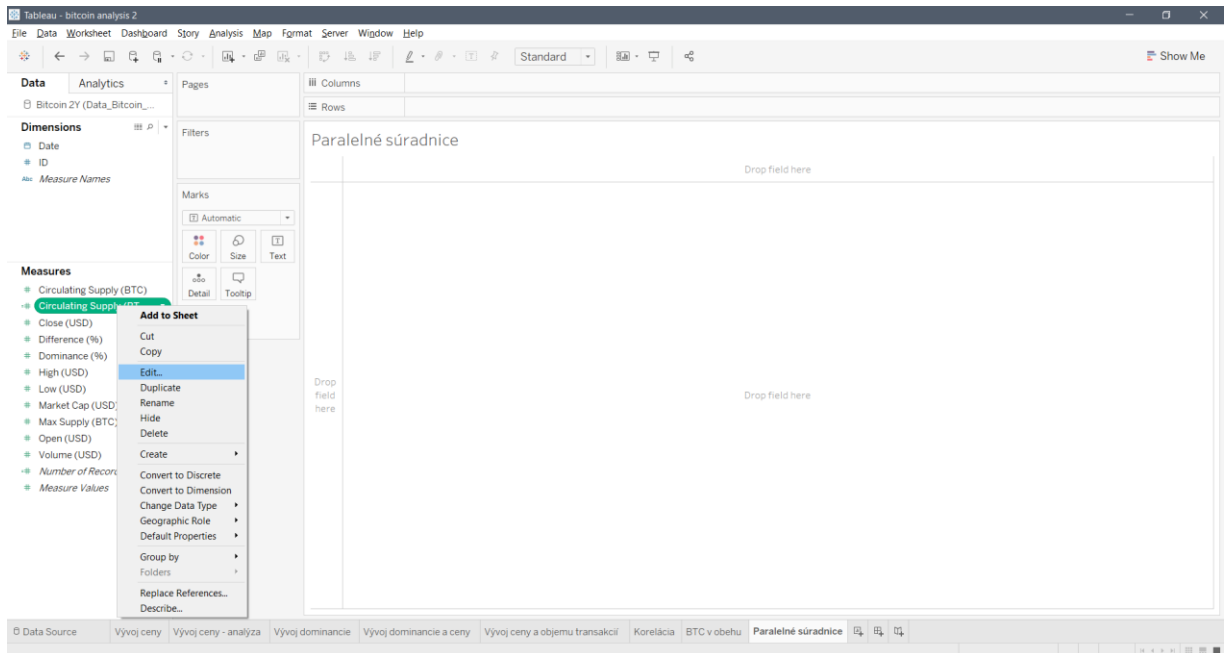
Obrázkový postup tvorby paralelnej súradnicovej sústavy



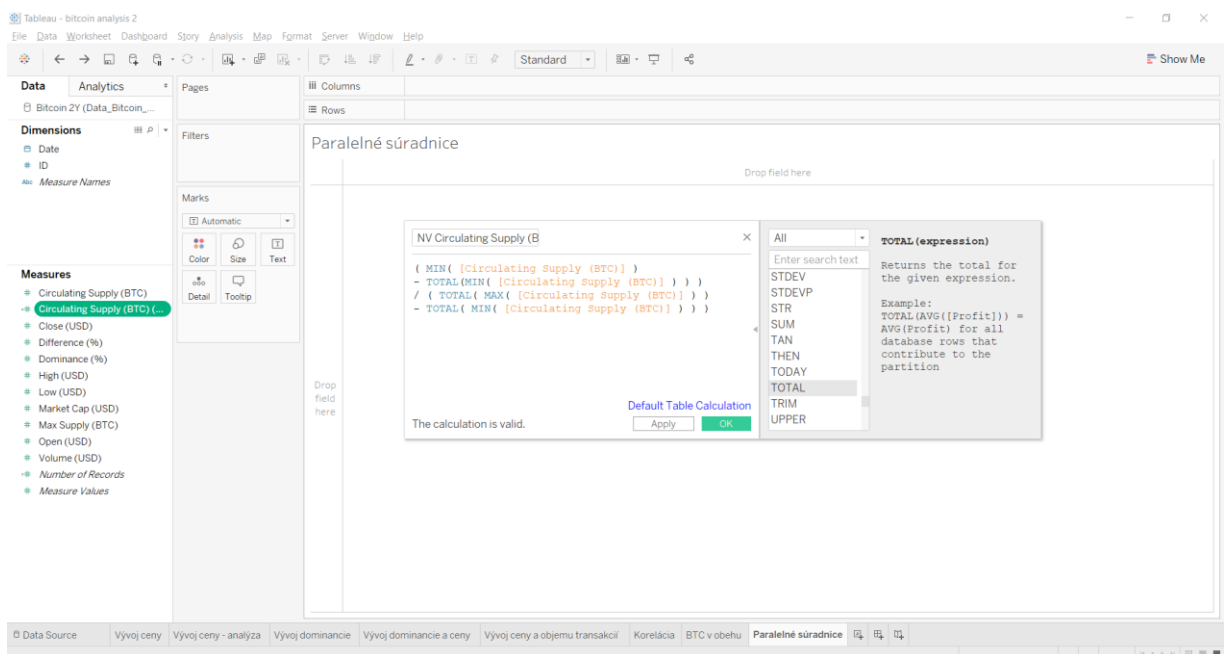
Obr. B1 - Vytvorenie novej záložky v Tableau



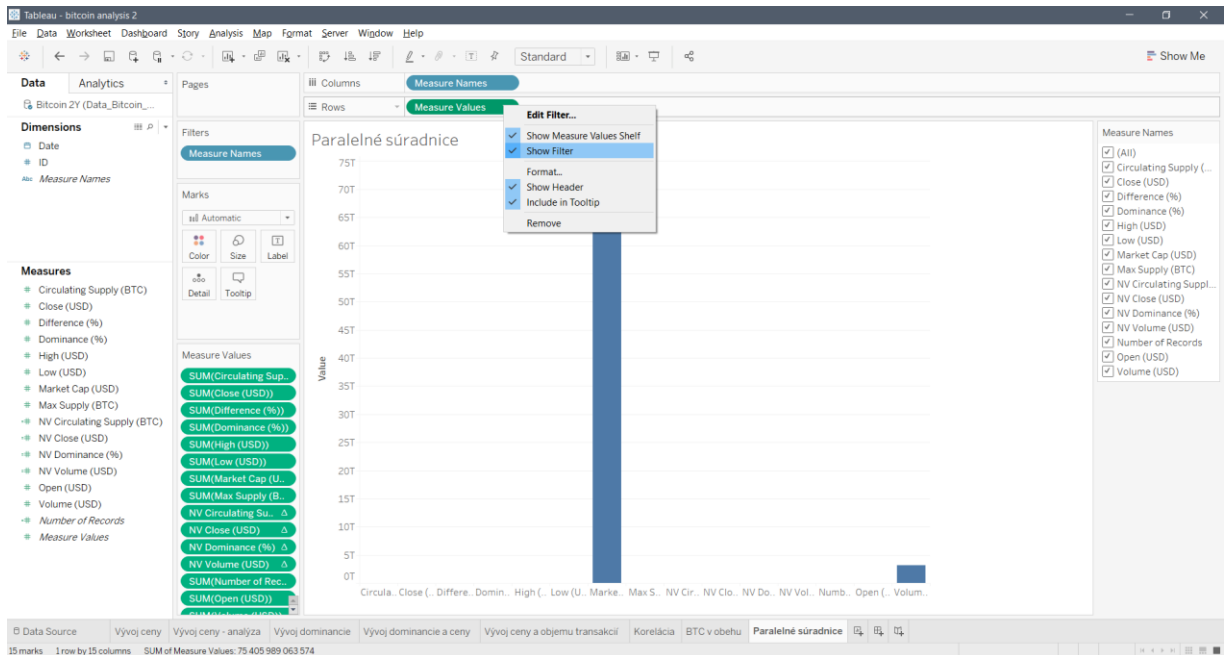
Obr. B2 - Duplikovanie parametrov meraní



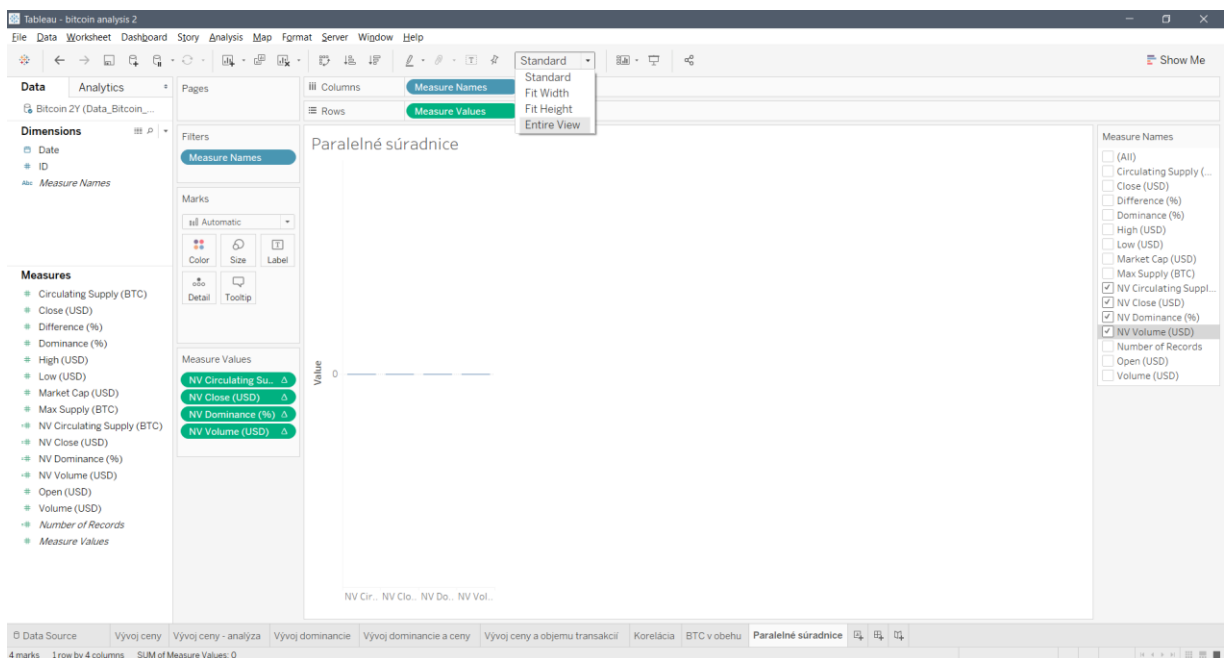
Obr. B3 - Otvorenie obrazovky na úpravu vzorca parametra merania



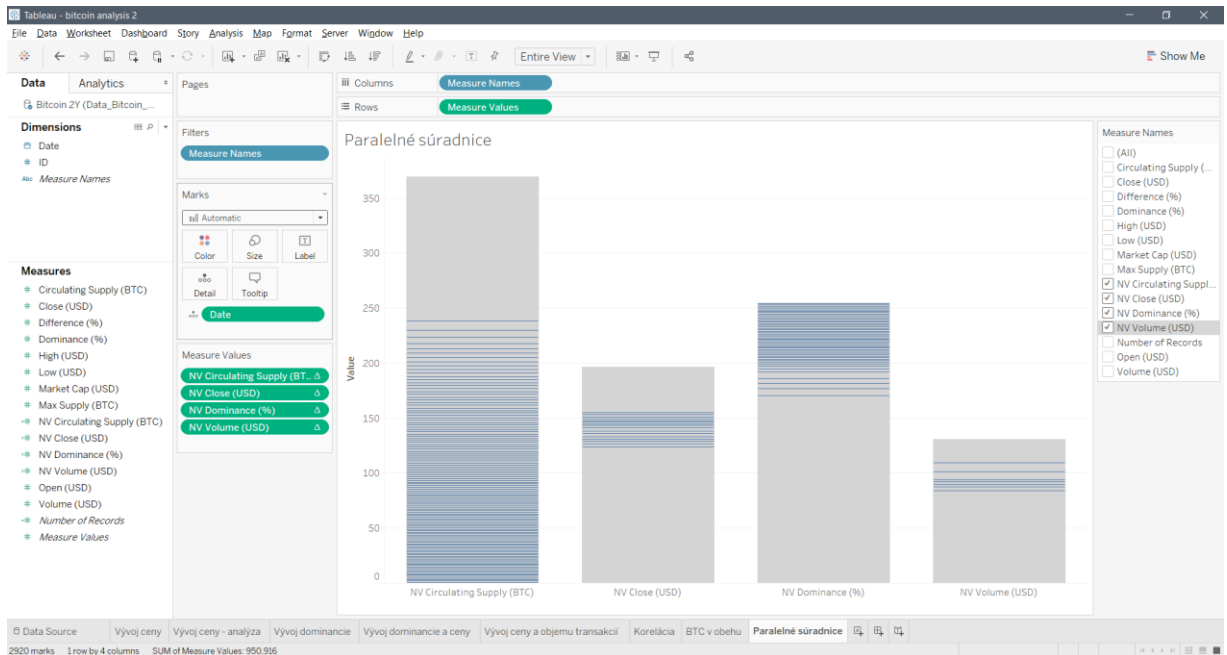
Obr. B4 - Zapísanie normalizačného vzorca do neagregovaného parametra merania



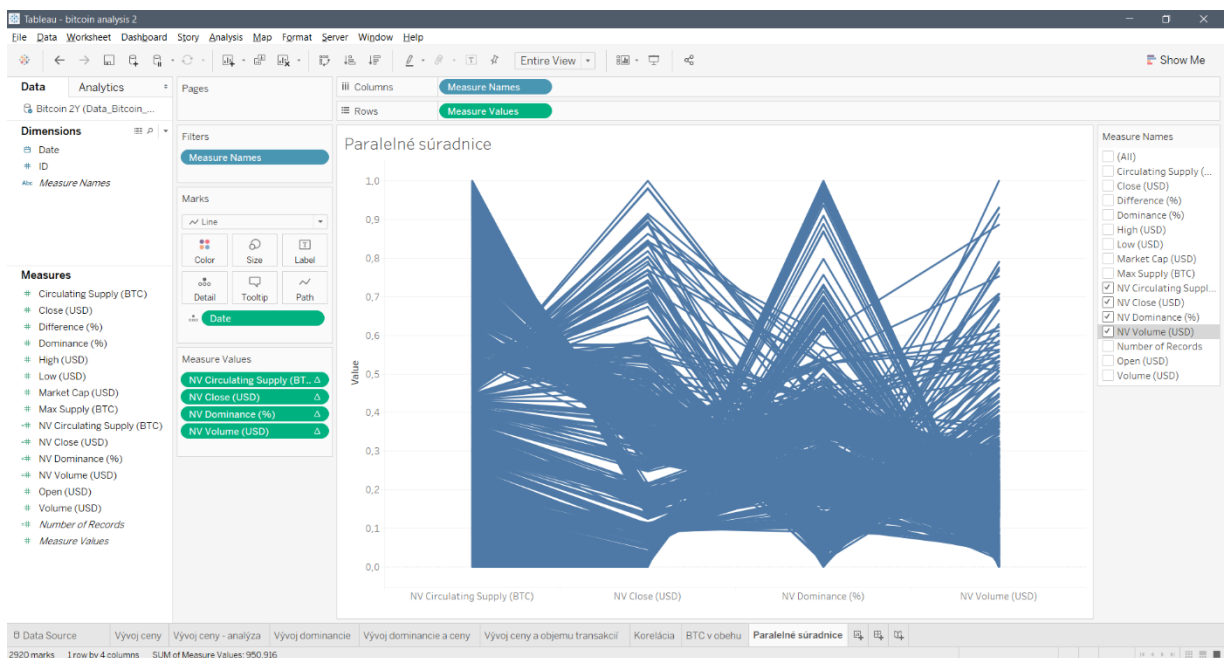
Obr. B5 - Zobrazenie filtra pre "Measure Names"



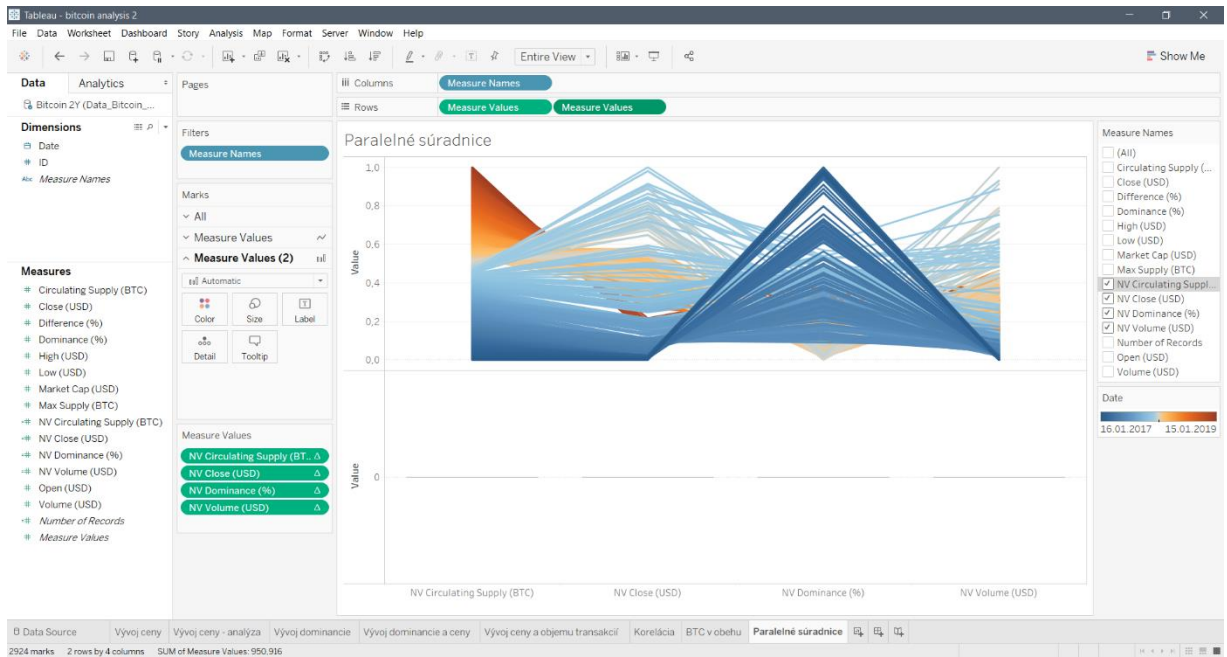
Obr. B6 - Výber položiek s normalizovanou škálou hodnôt



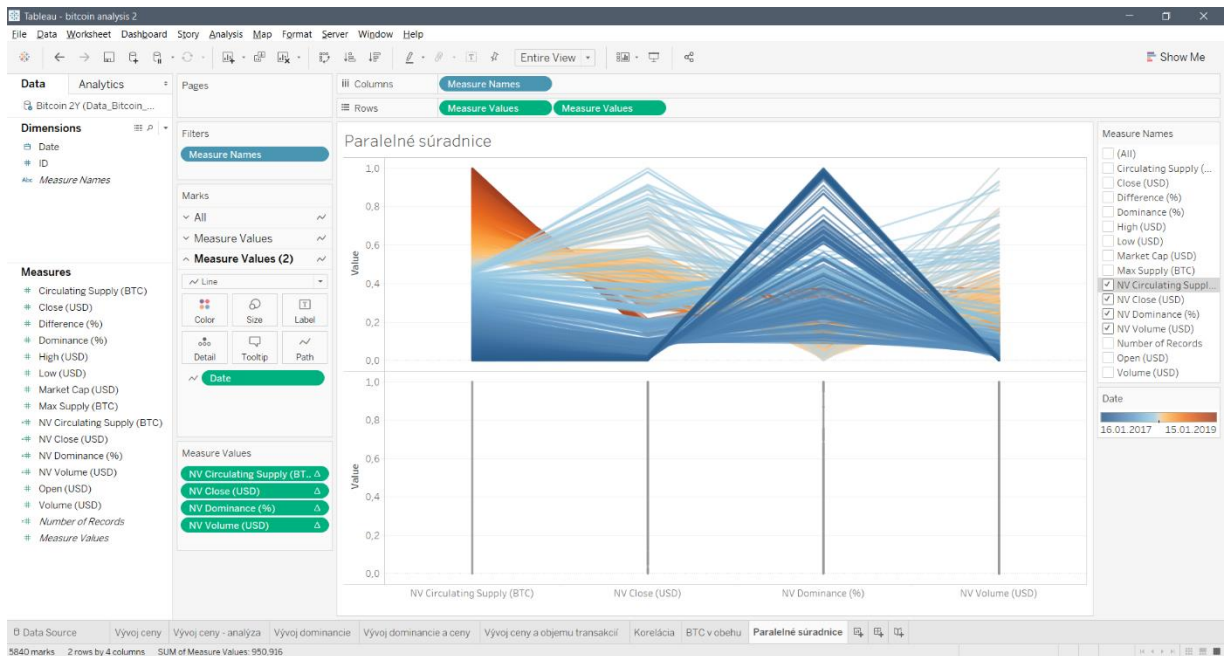
Obr. B7 - Pridanie rozmeru (dátumovej zložky) do vizualizácie



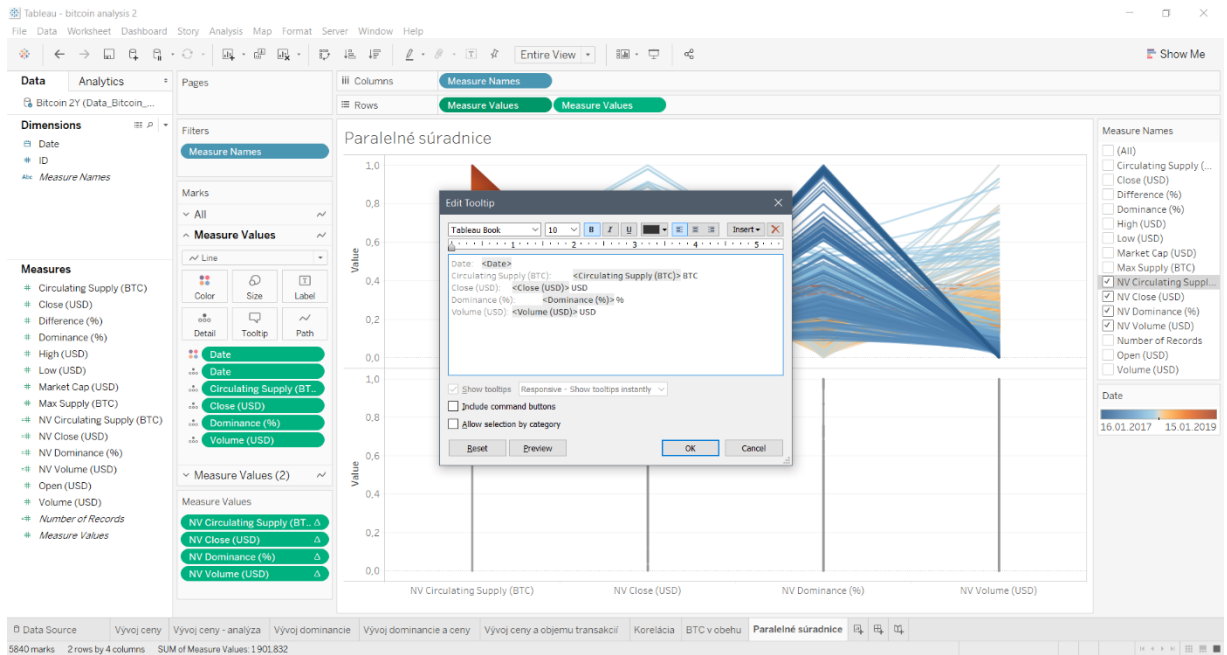
Obr. B8 - Zmena zobrazenia stĺpcov na lomené čiary v časti "Marks"



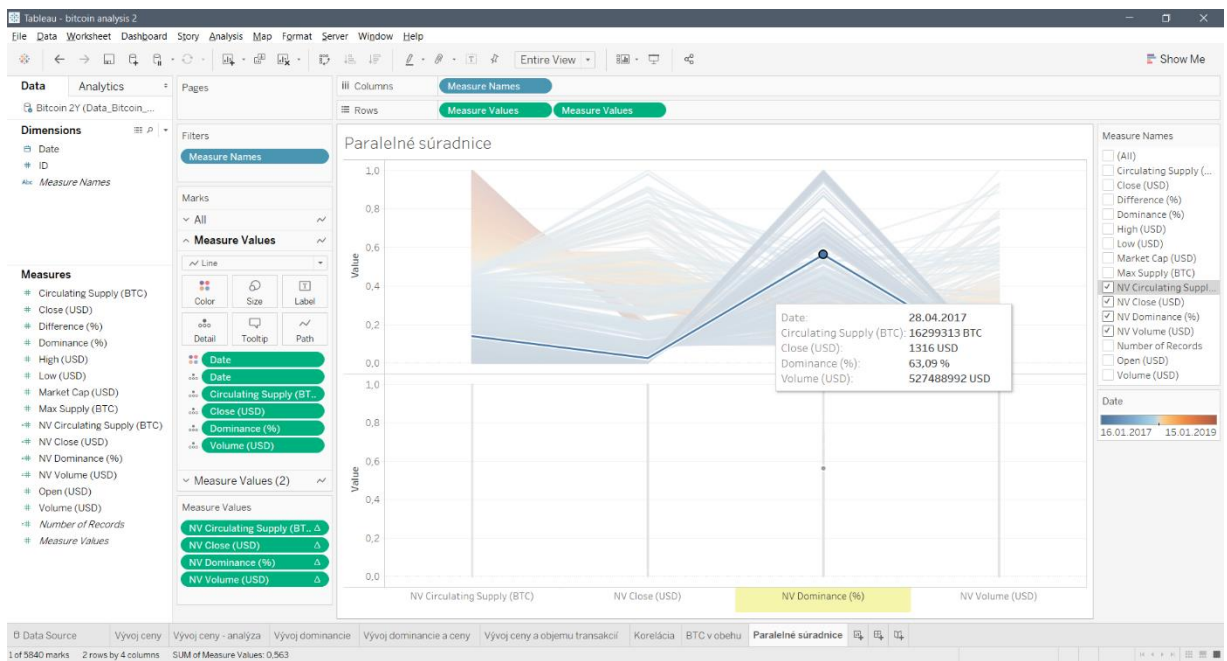
Obr. B11 - Pridanie základu pre paralelné osi do grafu



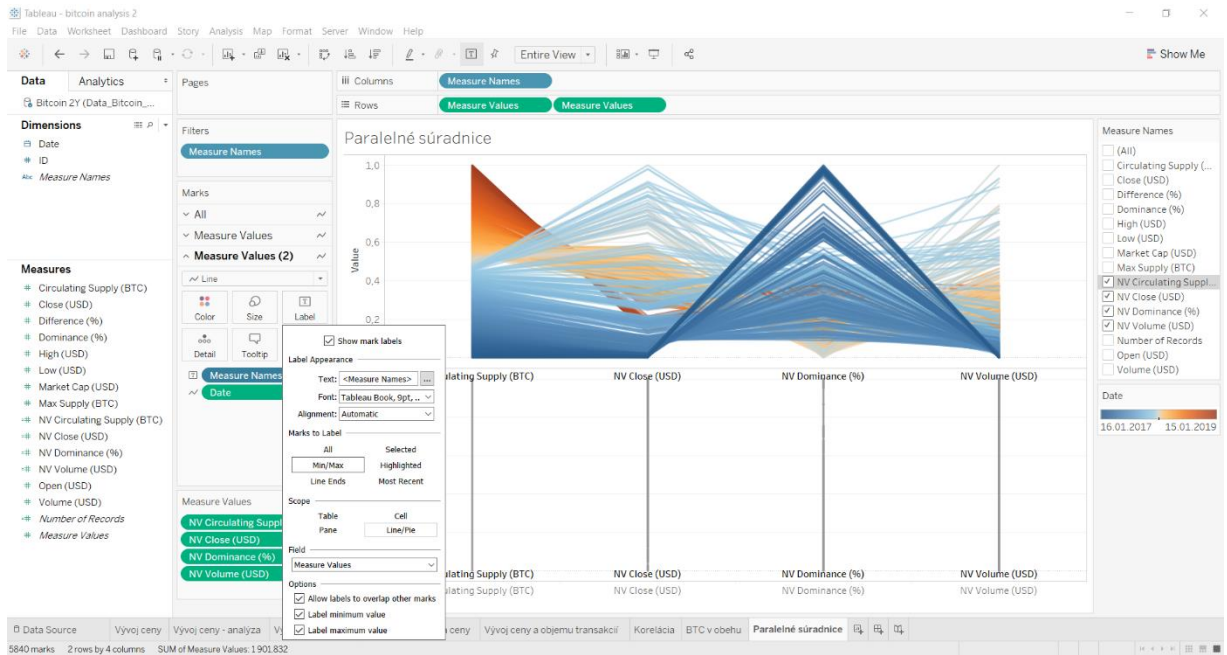
Obr. B12 - Vytvorenie paralelných priamok vďaka dátumovej dimenzii



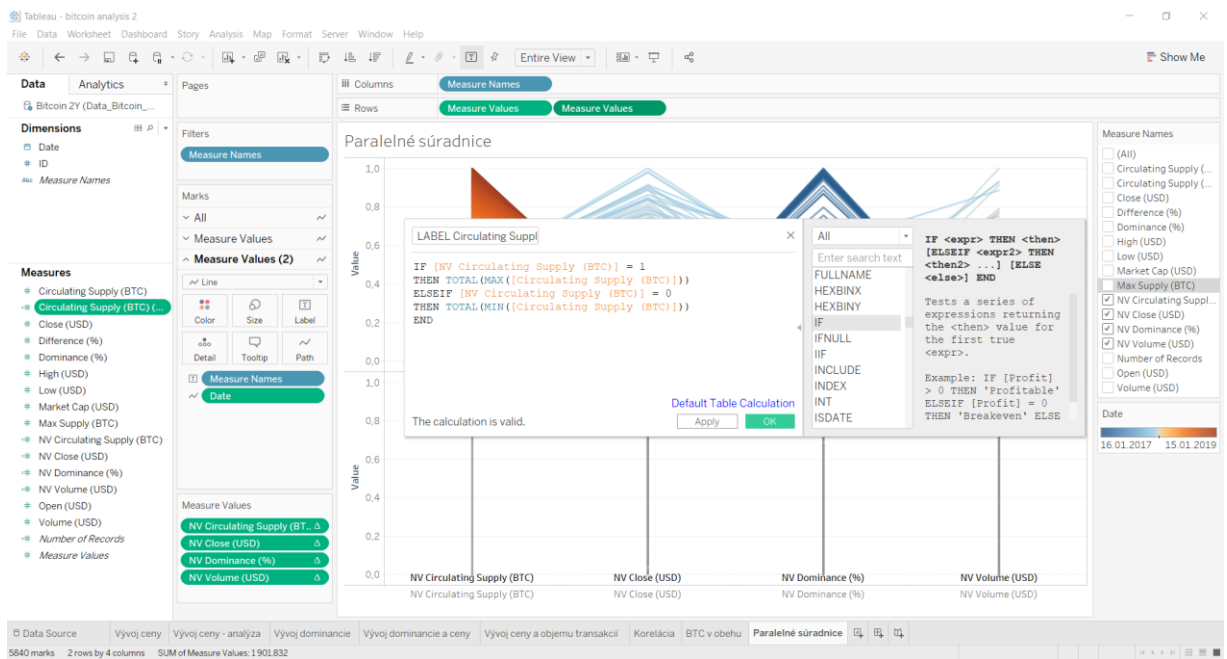
Obr. B13 - Úprava informačného štítka



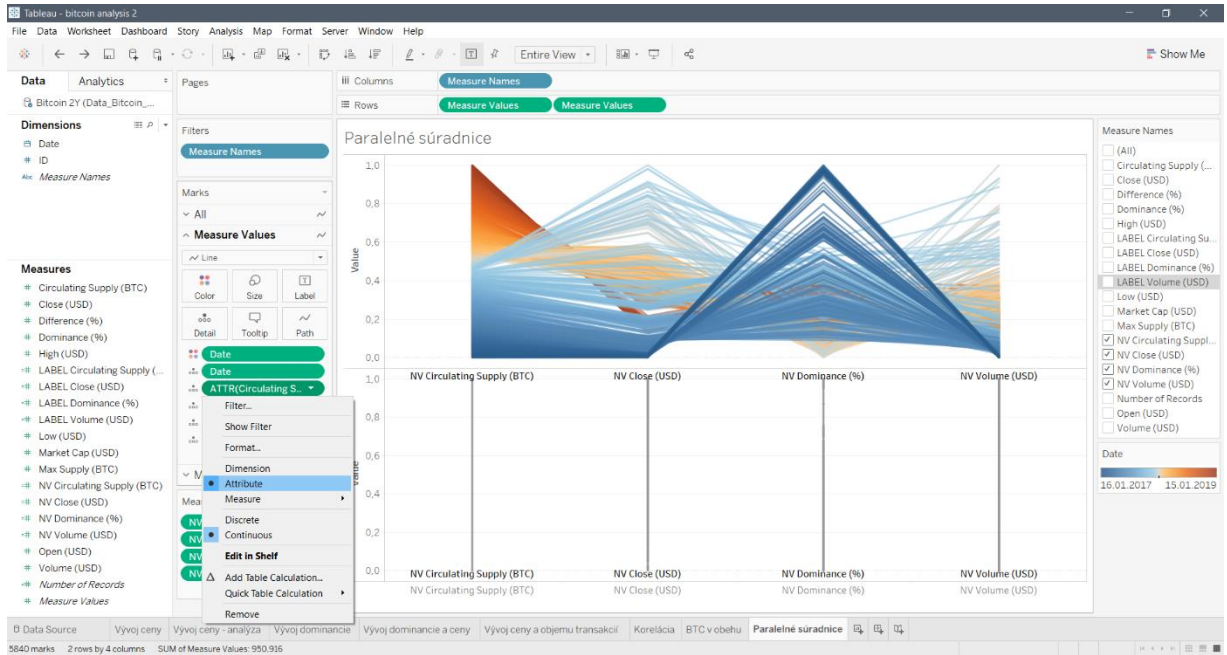
Obr. B14 - Príklad zobrazenia doplňujúcich informácií o konkrétnej lomenej čiare



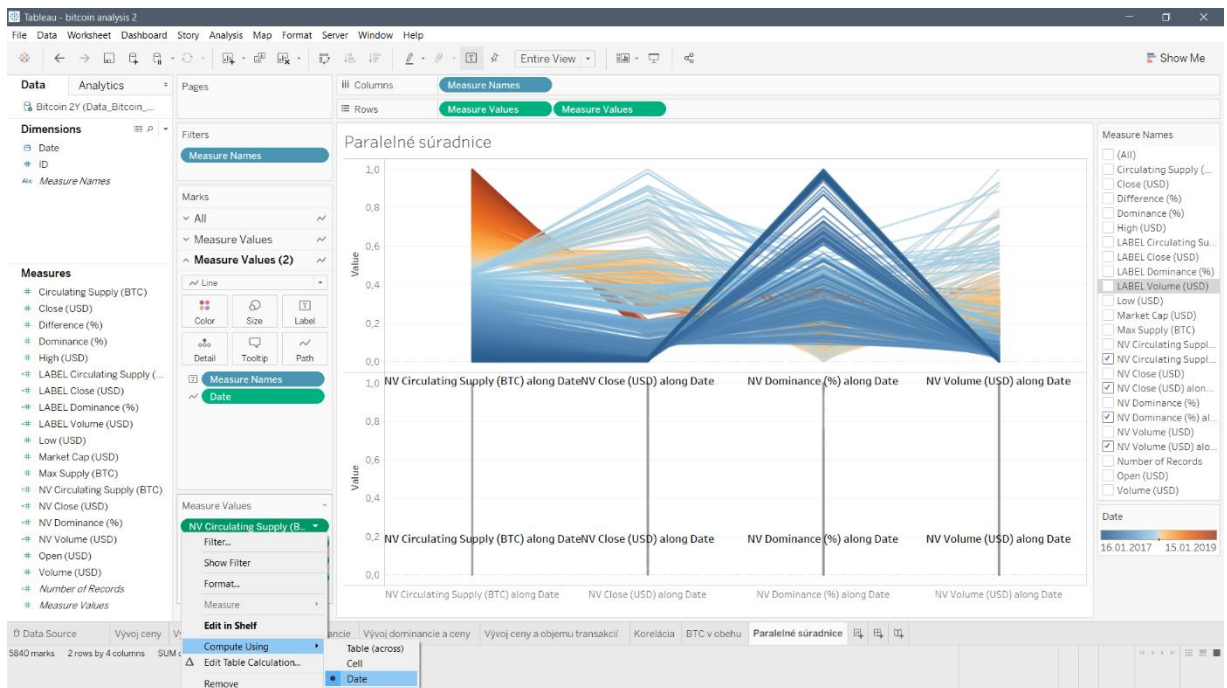
Obr. B15 - Pridanie názvov k jednotlivým paralelným osiam



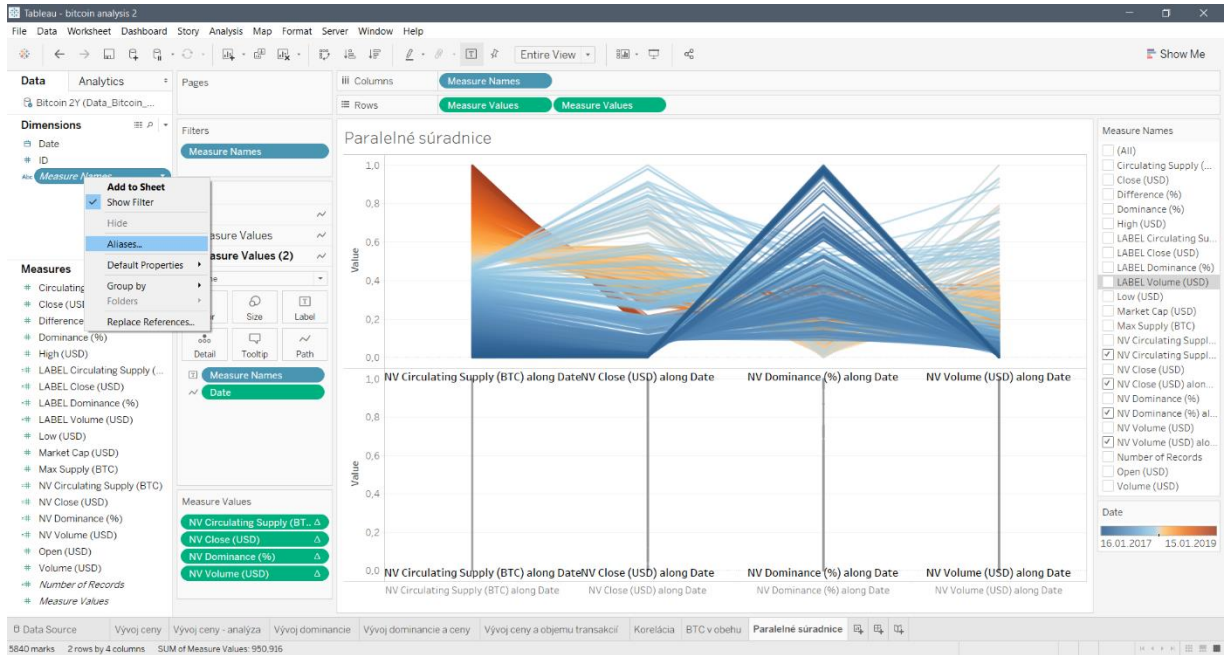
Obr. B16 - Vytvorenie novej premennej s podmienkou na zobrazovanie extrémov



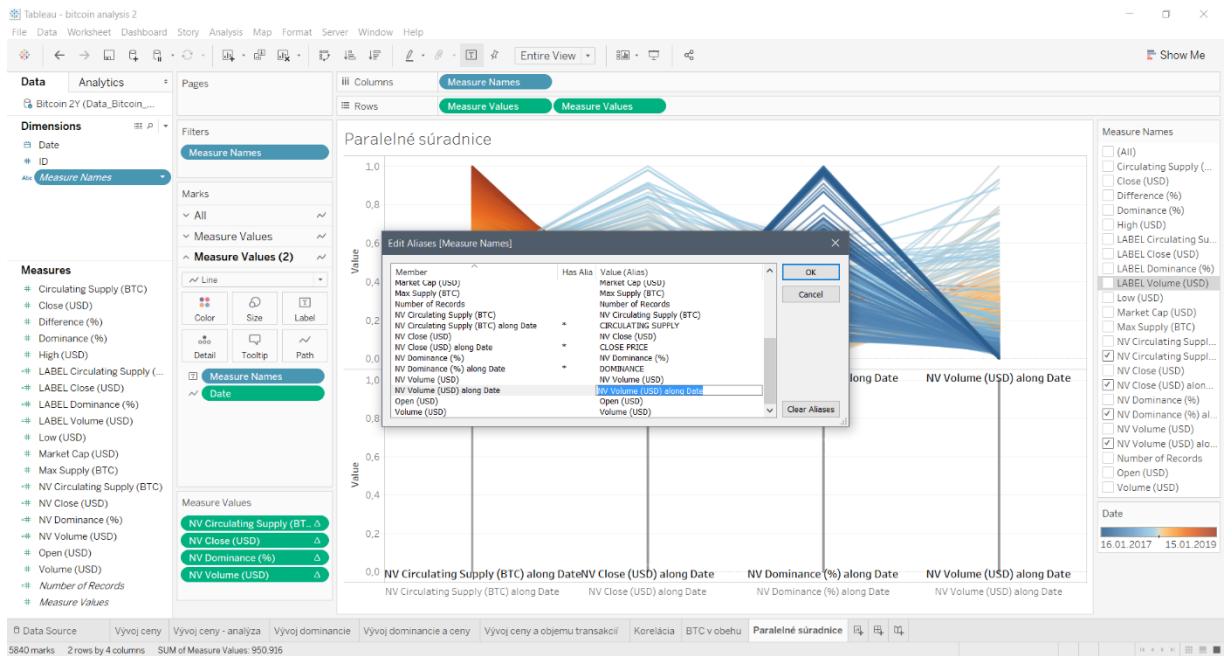
Obr. B17 - Zmena parametrov meraní na atribúty



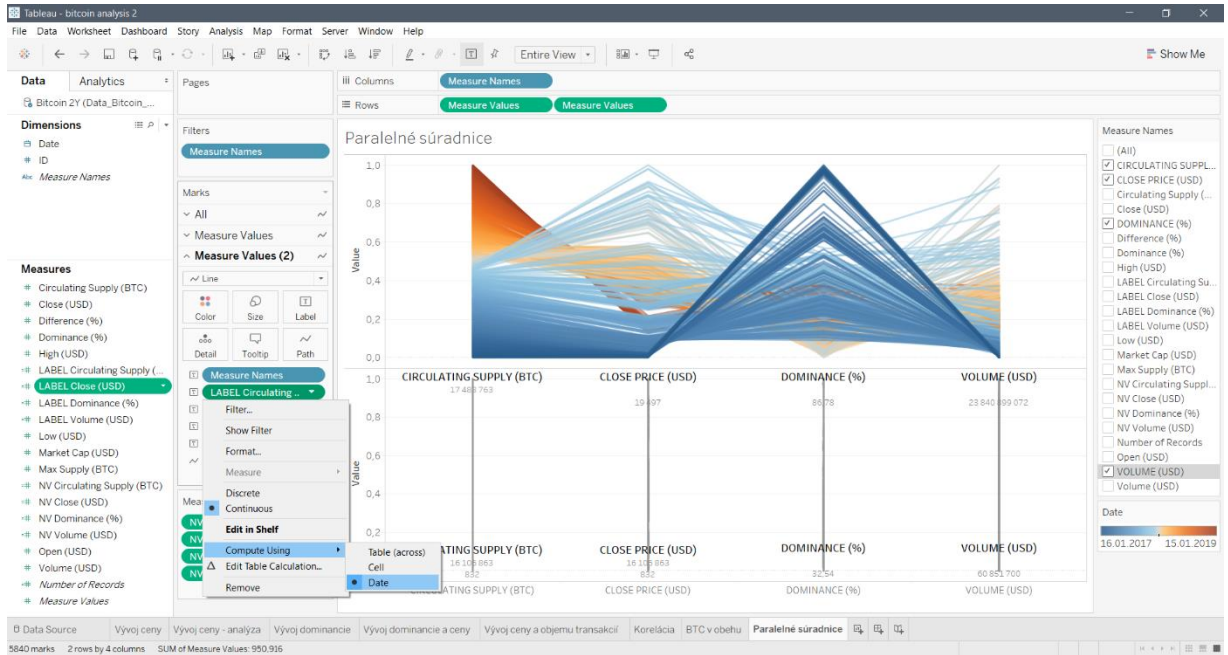
Obr. B18 - Výpočet vzorcov v parametroch cez dimenzionálnu zložku



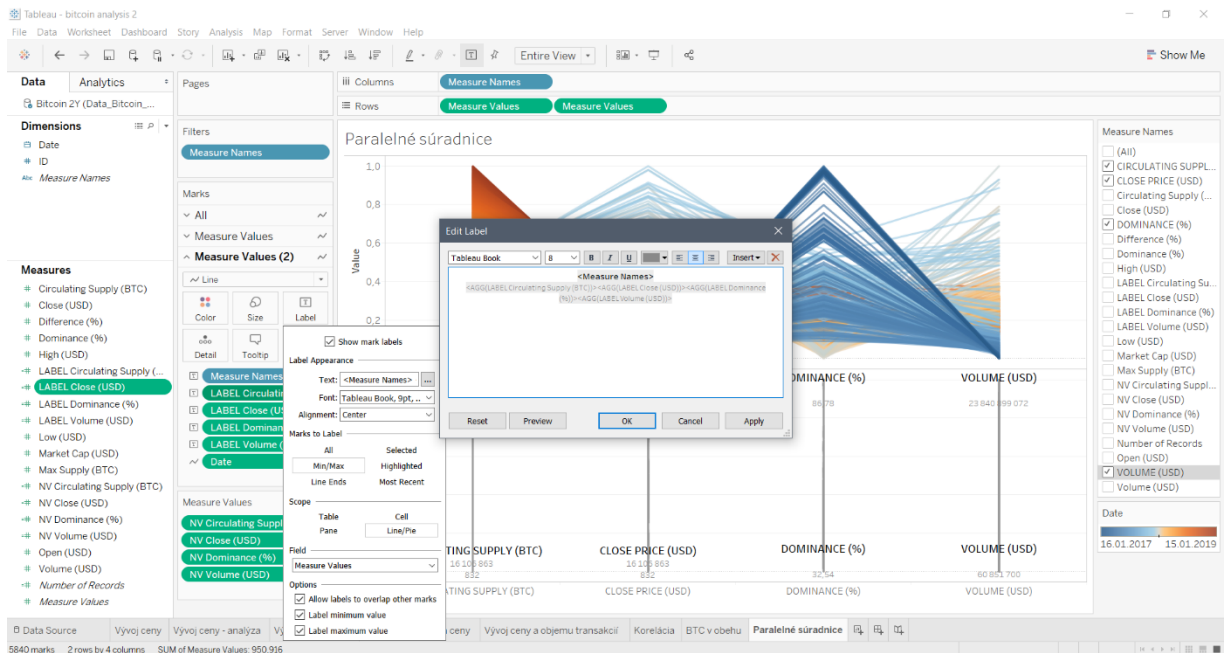
Obr. B19 - Výber možnosti na premenovanie premenných používaných vo vizualizácii



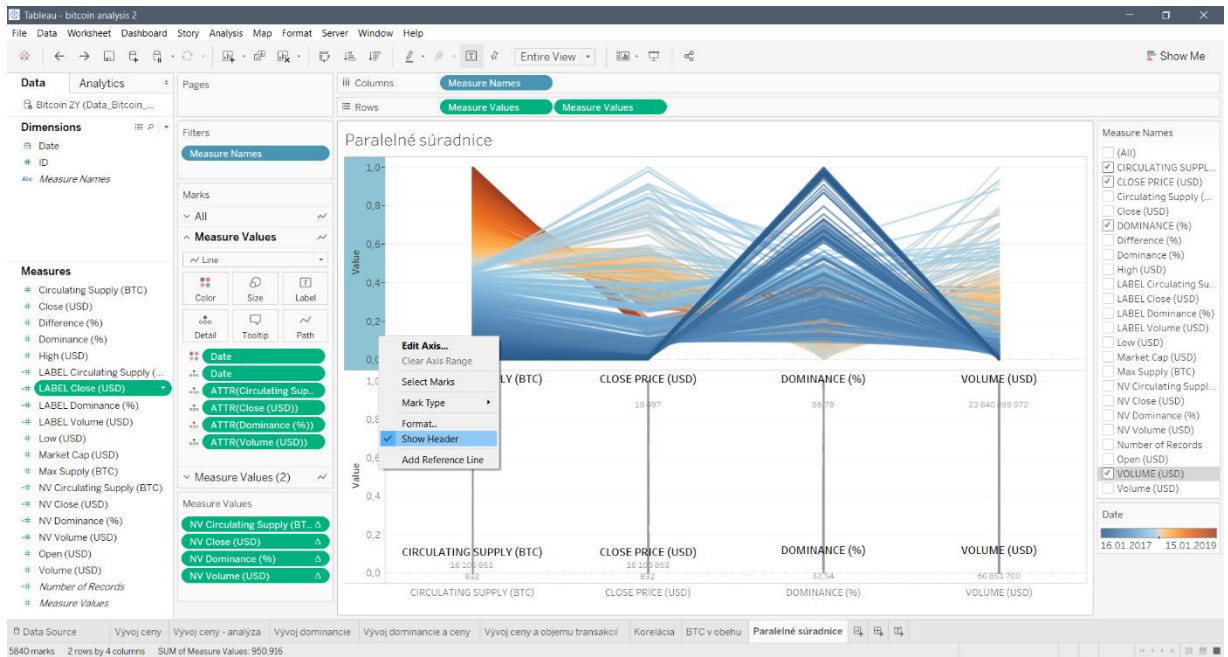
Obr. B20 - Premenovanie premenných používaných vo vizualizácii



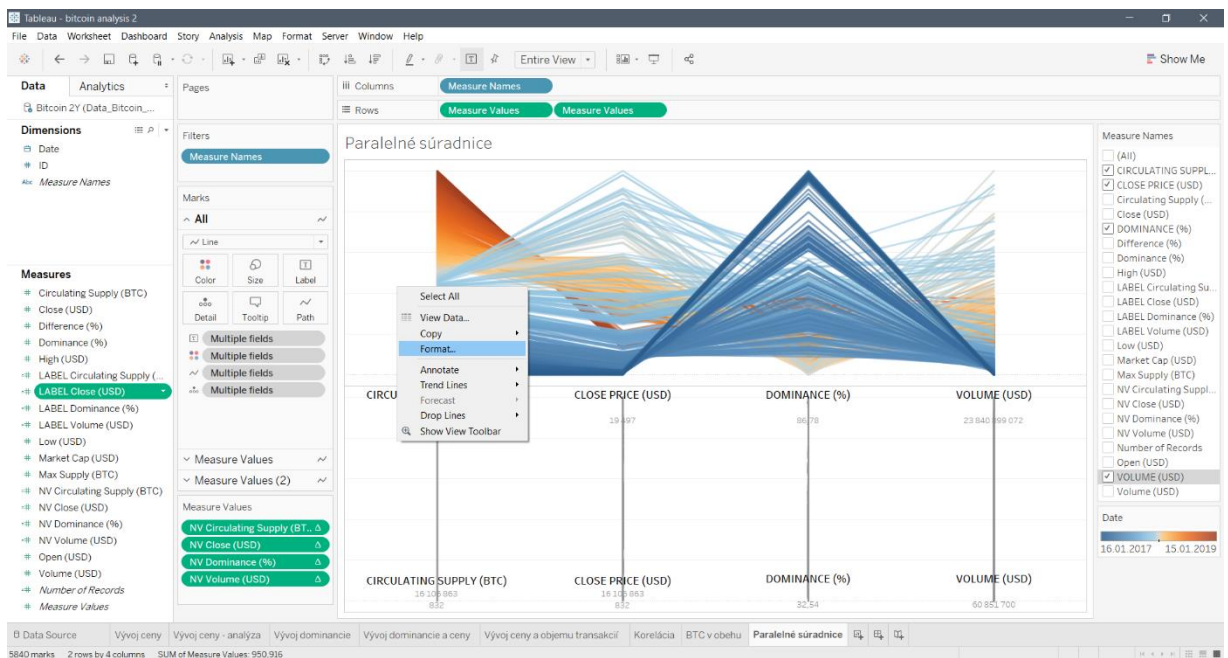
Obr. B21 - Nastavenie výpočtu popisujúcich parametrov cez dimenzionálnu zložku



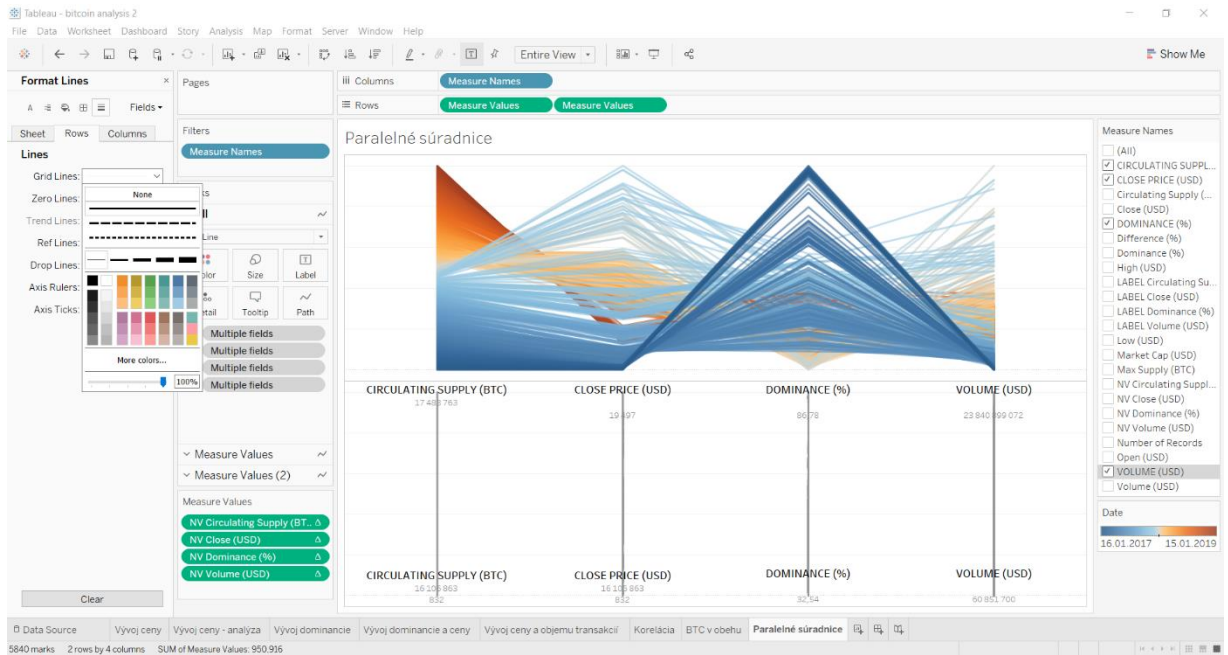
Obr. B22 - Zmena formátovania textu pri paralelných osiach v časti "Label"



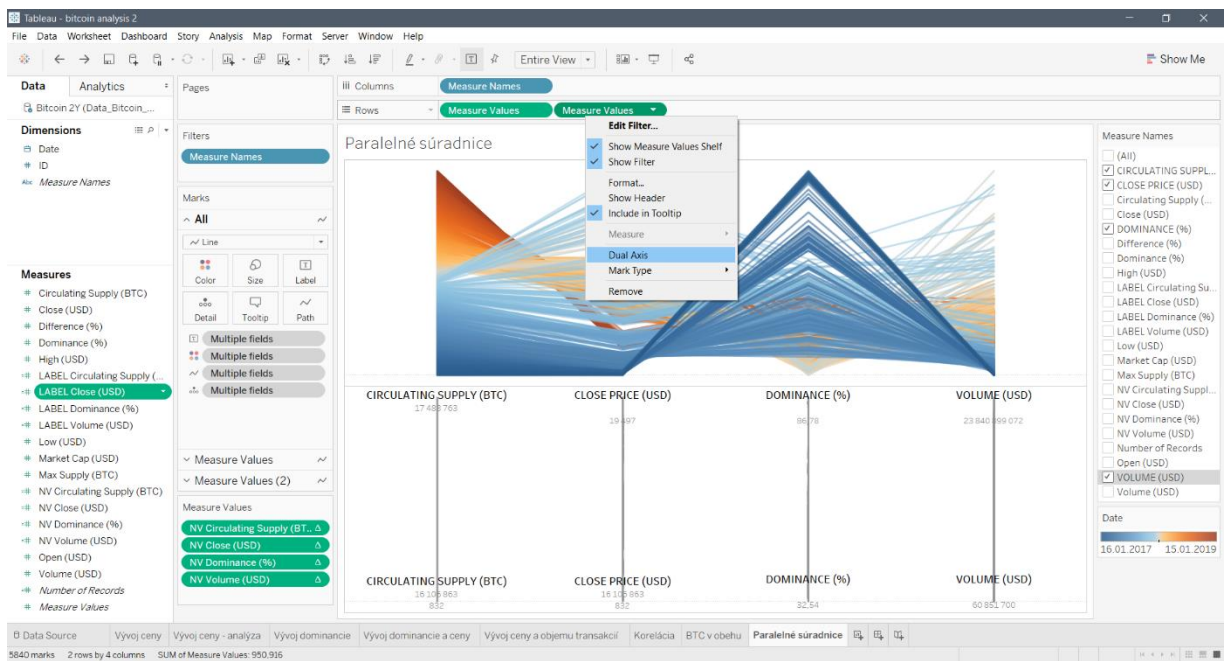
Obr. B23 - Zobrazenie alebo skrytie predvolených hlavičiek súradnicových osí



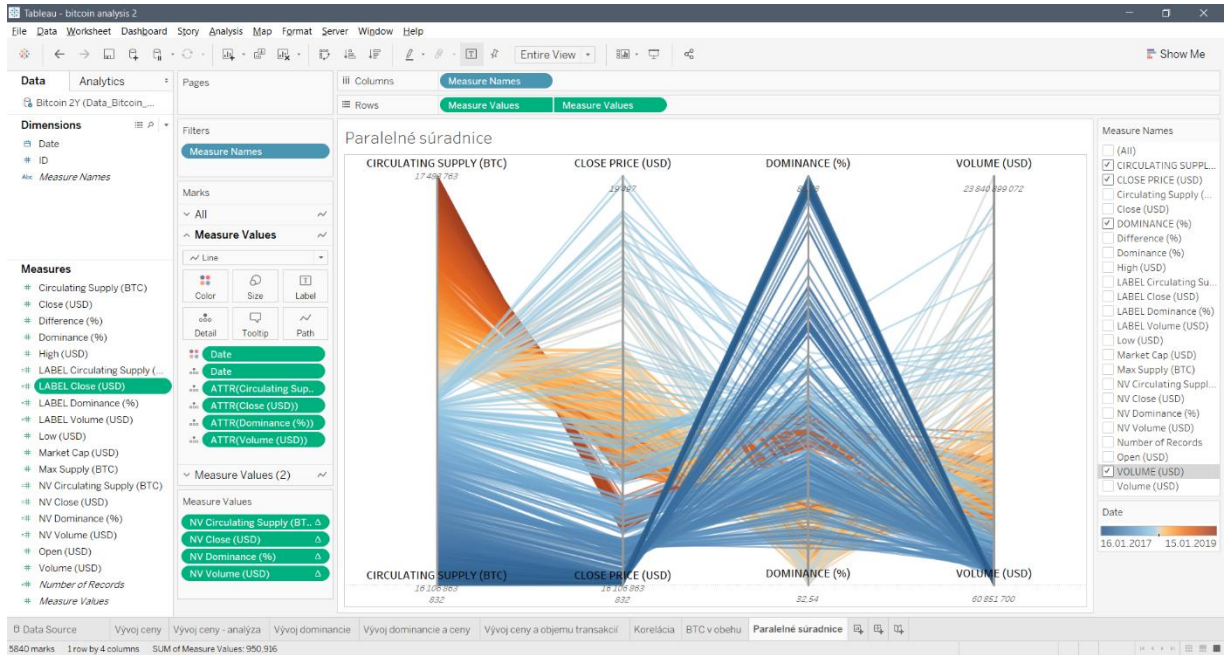
Obr. B24 - Otvorenie panelu s formátovaním



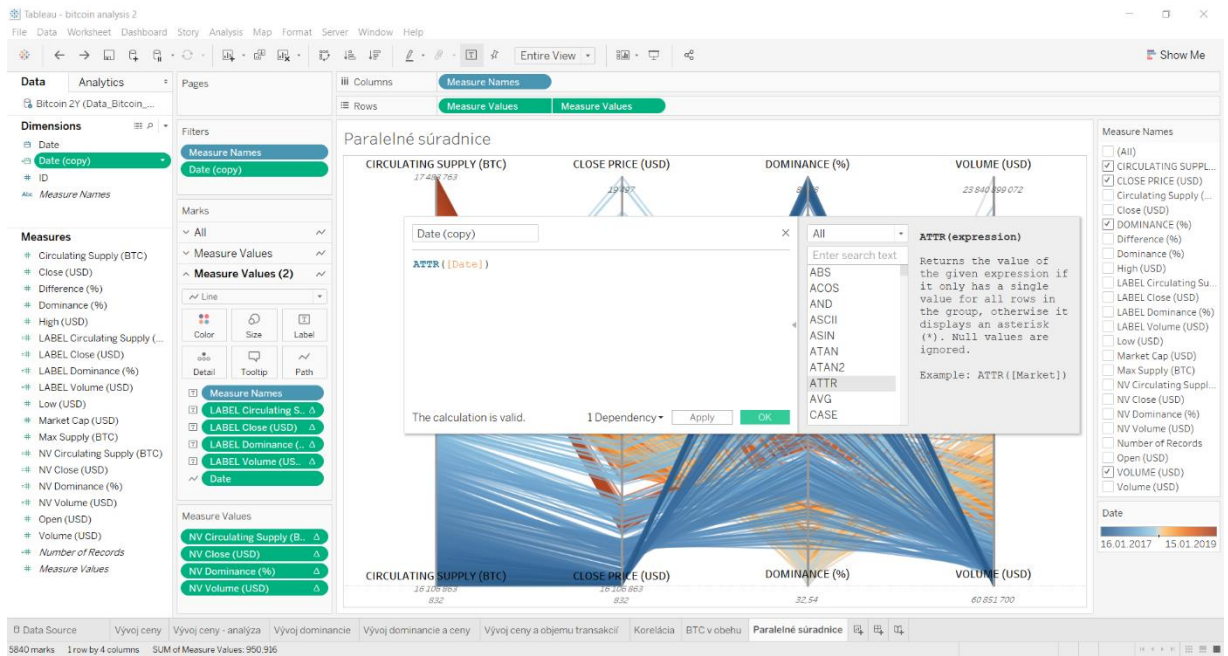
Obr. B25 - Zmena vodiacich čiar vo vizualizácii



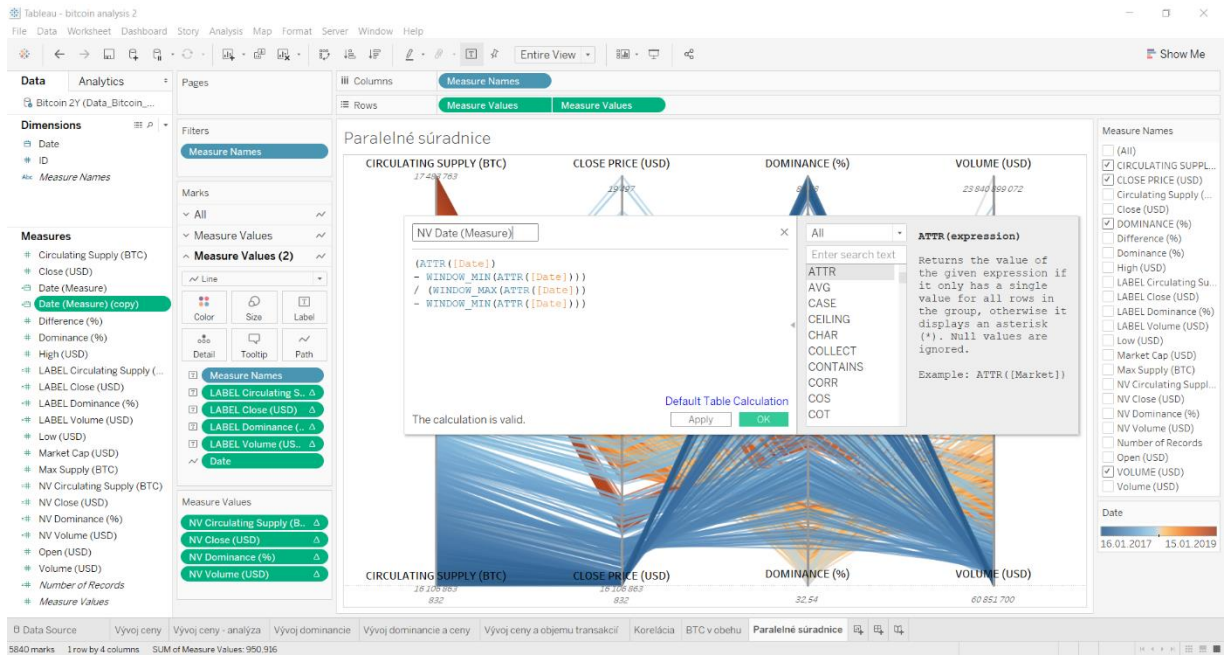
Obr. B26 - Umiestnenie dvoch grafov na jednu súradnicovú os



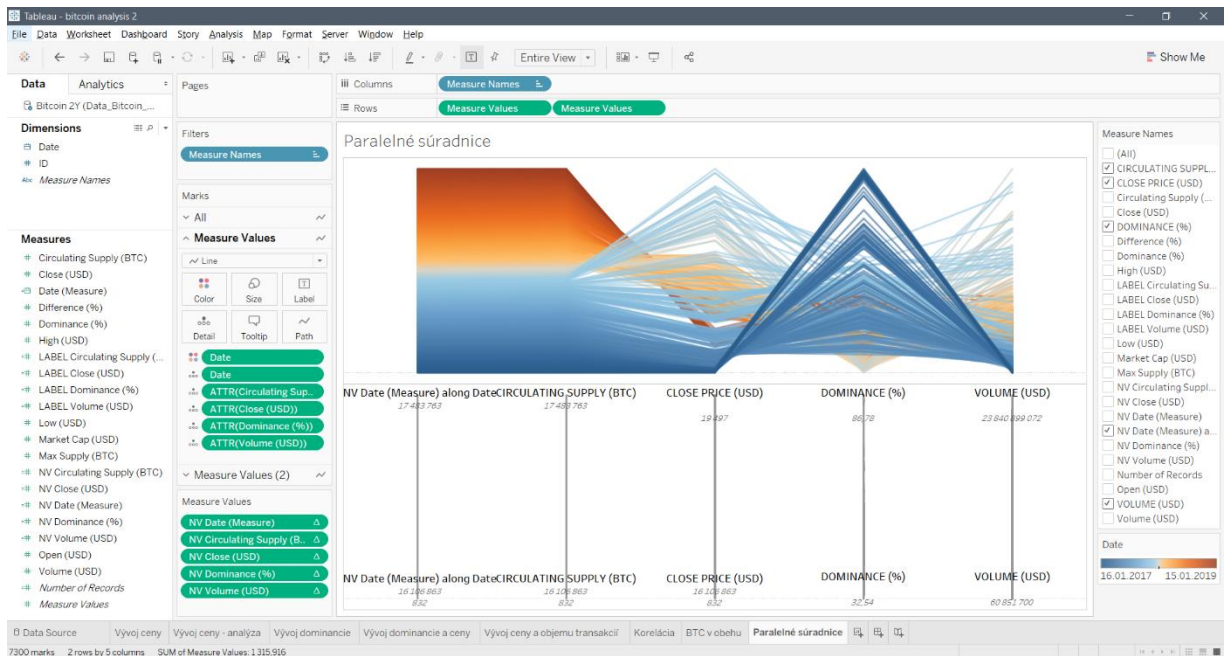
Obr. B27 - Výsledná vizualizácia v Tableau bez osi s dátumovou informáciou



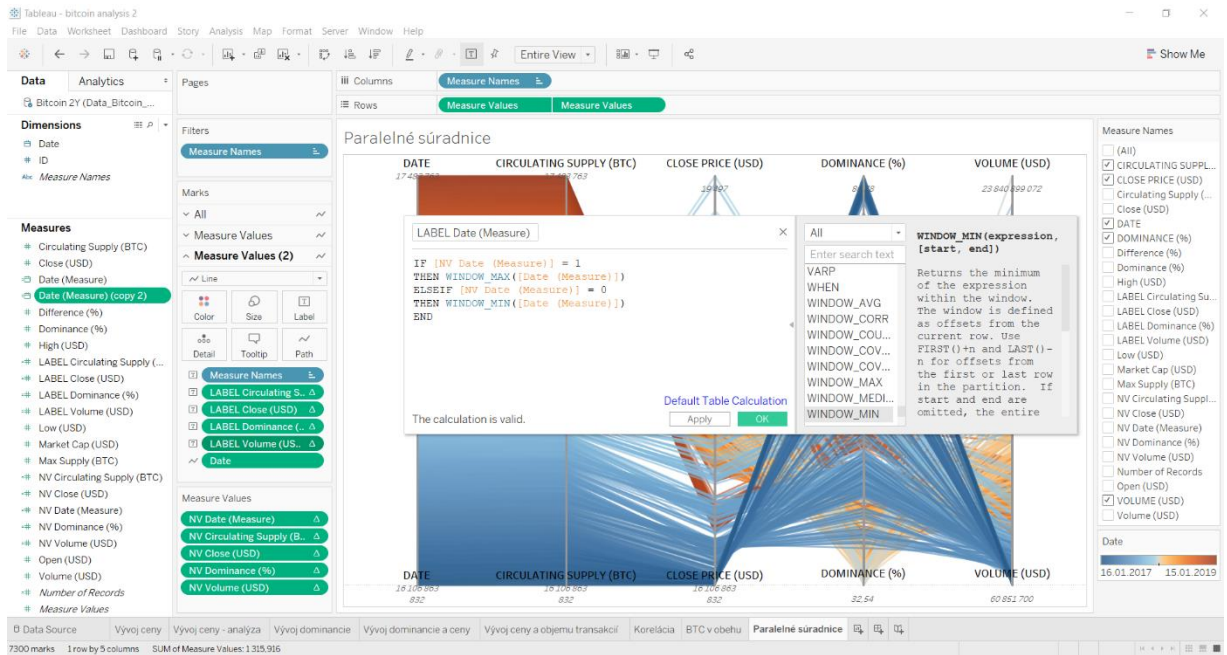
Obr. B28 - Upravený vzorec duplikovanej dátumovej zložky



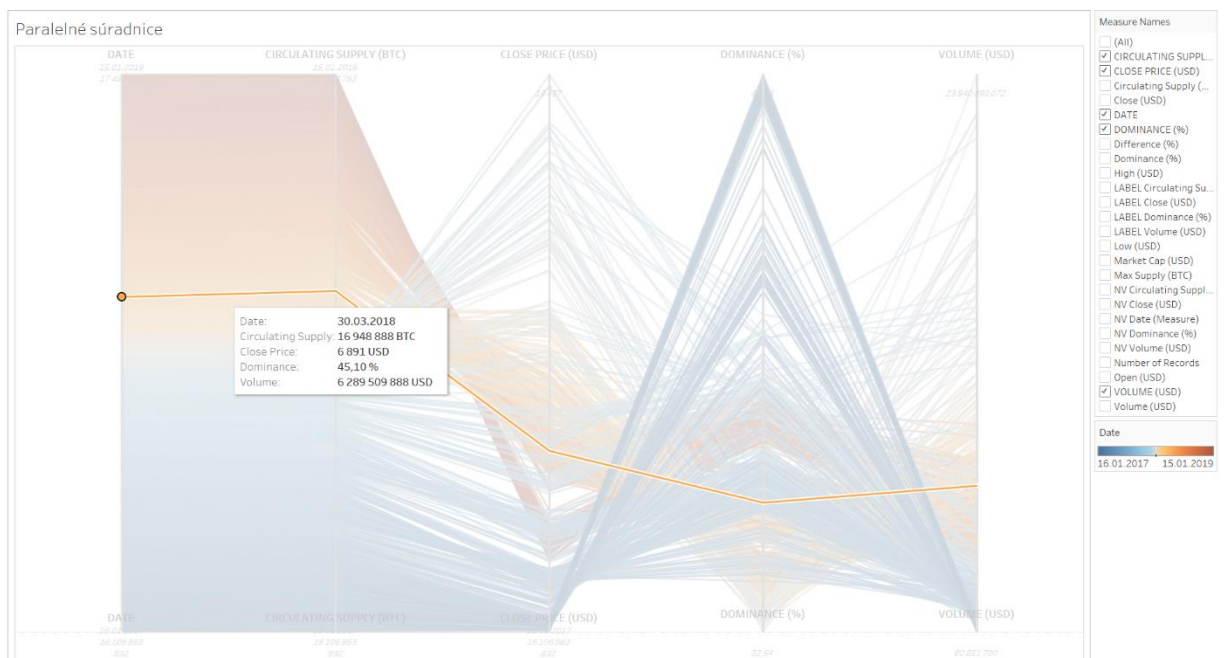
Obr. B29 - Normalizovanie dátumovej zložky pre potreby vytvorenia ďalšej paralelnej osi



Obr. B30 - Pridanie dátumového parametra do vizualizácie



Obr. B31 - Vzorec na zobrazovanie minimálnej a maximálnej hodnoty dátumovej premennej



Obr. B32 - Príklad vybraného záznamu vo finálnej vizualizácii