

Univerzita Pardubice

Fakulta ekonomicko-správní

Predikce volatility na kapitálových trzích pomocí analýzy sentimentu

Bc. Kateřina Šindelková

Diplomová práce

2016

ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: Bc. Kateřina Šindelková
Osobní číslo: E15915
Studijní program: N6208 Ekonomika a management
Studijní obor: Ekonomika a management podniku
Název tématu: Predikce volatility na kapitálových trzích pomocí analýzy sentimentu
Zadávající katedra: Ústav podnikové ekonomiky a managementu

Zásady pro vypracování:

Cílem práce je vysvětlit a blíže popsat problematiku volatility na kapitálových trzích. Dále se pokusit pomocí analýzy sentimentu predikovat volatilitu vybraných akcií firem obchodujících na zahraničním trhu. Analýza sentimentu bude provedena na základě dat z amerických akciových trhů s využitím regresní analýzy. Součástí práce bude také interpretace a vizualizace výsledků.

Osnova:

- Volatilita kapitálových trhů.
- Analýza sentimentu na kapitálových trzích.
- Sběr a předzpracování dat.
- Aplikace vhodného postupu pro získání informací pro predikci volatility pomocí analýzy sentimentu.
- Interpretace a vizualizace získaných výsledků.

Rozsah grafických prací:

Rozsah pracovní zprávy: cca 60 stran

Forma zpracování diplomové práce: tištěná/elektronická

Seznam odborné literatury:


LIU, Bing a rafael SAX. Sentiment analysis and opinion mining. 2nd ed. San Rafael: Morgan, c2012, xiv, 165 s. Synthesis lectures on human language technologies, 16. ISBN 978-1-60845-884-4

NEFTCI, Salih N. Principles of financial engineering. 2nd ed. Amsterdam: Elsevier, c2008, xiii, 679 s. Academic press advanced finance series. ISBN 978-012-3735-744

VESELÁ, Jitka. Investování na kapitálových trzích. Vyd. 1. Praha: ASPI, 2007, 703 s. ISBN 978-80-7357-297-6

ZMEŠKAL, Zdeněk, Dana DLUHOŠOVÁ a Tomáš TICHÝ. Finanční modely: koncepty, metody, aplikace. 3., přeprac. a rozš. vyd. Praha: Ekopress, 2013, viii, 267 s. Synthesis lectures on human language technologies, 16. ISBN 978-80-86929-91-0

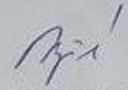
Vedoucí diplomové práce:


doc. Ing. Petr Hájek, Ph.D.

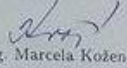
Ústav systémového inženýrství a informatiky

Datum zadání diplomové práce: 29. září 2015

Termín odevzdání diplomové práce: 30. dubna 2016


doc. Ing. Renáta Myšková, Ph.D.
děkanka

L.S.


doc. Ing. Marcela Kožená, Ph.D.
vedoucí ústavu

V Pardubicích dne 29. září 2015

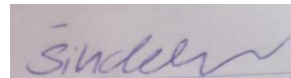
PROHLÁŠENÍ

Prohlašuji, že jsem tuto práci vypracovala samostatně. Veškeré literární prameny a informace, které jsem v práci využila, jsou uvedeny v seznamu použité literatury.

Byla jsem seznámena s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorský zákon, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako Školního díla podle § 60 odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Souhlasím s prezenčním zpřístupněním své práce v Univerzitní knihovně Univerzity Pardubice.

V Pardubicích dne 30.6.2016

A rectangular box containing a handwritten signature in blue ink, which appears to be 'Šindelová'.

PODĚKOVÁNÍ:

Tímto bych ráda poděkovala svému vedoucímu práce doc. Ing. Petru Hájkovi, Ph.D. za jeho odbornou pomoc, cenné rady a poskytnuté materiály, které mi pomohly při zpracování diplomové práce.

Děkuji též své rodině a svým blízkým, kteří mi po dobu psaní této práce poskytovali podporu.

ANOTACE

Tato diplomová práce je zaměřena na analýzu sentimentu v oblasti financí. První část je zaměřena na trh cenných papírů, volatilitu, analýzu sentimentu a psychologii investování. Po teoretickém vysvětlení zmíněných témat je proveden sběr a zpracování dat pro analýzu sentimentu v internetových článcích týkajících se vybraných podniků. Následně je pomocí regresní analýzy posouzena závislost mezi sentimentem a vývojem volatility na kapitálových trzích.

KLÍČOVÁ SLOVA

analýza sentimentu, názory, slovník sentimentu, regresní analýza, volatilita, psychologie

TITLE

Prediction of volatility in the capital markets by sentiment analysis

ANNOTATION

This thesis is focused on the sentiment analysis especially in finance. The theoretical part is focused on securities market, volatility, sentiment analysis and psychology of investing. After the theoretical explanation of mentioned topics, data are collected and processed for the sentiment analysis in the internet articles related to the selected companies. Subsequently the dependence of the sentiment and the volatility development on the capital markets is evaluated using the regression analysis.

KEYWORDS

Sentiment analysis, Opinions, Sentiment Lexicon, Regression analysis, volatility, psychology

OBSAH

ÚVOD.....	11
1 TRH CENNÝCH PAPÍRŮ	13
1.2 Struktura trhu cenných papírů	13
1.3 Mimoburzovní trh Nasdaq.....	14
1.4 Fundamentální vlivy na vývoj ceny akcií	15
1.5 Akciové indexy	15
1.6 Riziko	17
1.6.1 Nediverzifikovatelné riziko.....	17
1.6.2 Diverzifikovatelné riziko a analýza portfolia.....	18
2 VOLATILITA.....	20
2.1 Model oceňování kapitálových aktiv (CAPM).....	21
2.2 Druhy volatility:.....	22
2.3 Riziko volatility	23
2.4 EWMA model.....	23
2.5 Indexy volatility	24
3 PSYCHOLOGIE INVESTOVÁNÍ	27
3.1 Interní chyby investora.....	27
3.2. Externí chyby	29
3.3 Typy investorů.....	30
4 ANALÝZA SENTIMENTU.....	31
4.1 Postup při analýze sentimentu	33
4.2 Metody analýzy sentimentu.....	36
4.3 Slovníky	36
5 SBĚR A PŘÍPRAVA DAT PRO ANALÝZU.....	37
5.1 Vyhodnocení sentimentu	37
5.2 Ukazatele pro predikci volatility.....	40
5.2.1 Tržní kapitalizace	40
5.2.2 P/E	41
5.2.3 Volume (objem obchodů)	41
6 PREDIKCE VOLATILITY POMOCÍ REGRESNÍHO MODELU	42
6.1 Vícerozměrná lineární regrese.....	44
6.2 Výsledky vícerozměrné regrese.....	45

6.3 Ověření předpokladů modelu	48
6.4 Vizualizace predikované volatility	53
ZÁVĚR	57
SEZNAM POUŽITÉ LITERATURY	59
PŘÍLOHY	61

SEZNAM OBRÁZKŮ A TABULEK

Obrázek 1: Grafické znázornění přímky CP	22
Obrázek 2: Porovnání volatilit Apple a S&P 500	25
Obrázek 3: Porovnání volatilit Cisco a S&P 500	25
Obrázek 4: Porovnání volatilit Morgan Stanley a S&P 500	26
Obrázek 5: Porovnání volatilit Pepsi a S&P 500	26
Obrázek 6: Působení emocí při obchodování	29
Obrázek 7: Postup při analýze sentimentu	33
Obrázek 8: Ukázka prezentace analýzy sentimentu pomocí grafu	35
Obrázek 9: Vyhodnocení sentimentu firem	39
Obrázek 10: Metoda nejmenších čtverců	43
Obrázek 11: Normální p-graf reziduí 1. regrese	49
Obrázek 12: Normální p-graf reziduí 2. regrese	49
Obrázek 13: Histogram reziduí 1. regrese	50
Obrázek 14: Histogram reziduí 2. regrese	51
Obrázek 15: Bodový graf reziduí 1. regrese	52
Obrázek 16: Bodový graf reziduí 2. regrese	53
Obrázek 17: Porovnání volatilit model vs. realita Wells Fargo	54
Obrázek 18: Porovnání volatilit model vs. realita Berkshire Hathaway	54
Obrázek 19: Porovnání volatilit model vs. realita Morgan Stanley	55
Obrázek 20: Porovnání volatilit model vs. realita Ebay	55
Obrázek 21: Porovnání volatilit model vs. realita Visa	56
Tabulka 1: Nejčastější negativní slova z celé 10-K zprávy	36
Tabulka 2: Vyhodnocení sentimentu - pozitivní slova	38
Tabulka 3: Vyhodnocení sentimentu - negativní slova	38
Tabulka 4: Základní statistiky proměnné Typ burzy	40
Tabulka 5: Korelace nezávislých proměnných	45
Tabulka 6: Výsledky vícerozměrné regrese 1	46
Tabulka 7: Výsledky vícerozměrné regrese 2	47
Tabulka 8: Základní statistiky vybraných proměnných	48

SEZNAM ZKRATEK

CP	Cenné papíry
Nasdaq	National Association of Securities Dealers Automated Quotations
DJIA	Dow Jones Industrial Average
FTSE	Financial Times Stock Exchange
S&P 500	The Standard & Poor's 500
NYSE	The New York Stock Exchange
CAPM	Capital asset pricing model
SML	Security market line
CBOE	Chicago Board Options Exchange
VIX	Volatility index

ÚVOD

Analýza sentimentu je poměrně nová metoda analýzy textu, které slouží ke zjištění lidského sentimentu a názorů nejčastěji na internetu. Jelikož význam internetu za poslední dobu výrazně vzrostl, nelze přehlížet názory, které se na internetu začínají množit (ať už ve článkách, recenzích nebo komentářích). Konkrétně se může jednat o názory v podobě recenzí na zakoupený produkt, hodnocení kvality poskytovaných služeb nebo subjektivní názor na konkrétní podnik. Všechny zmíněné komentáře ovlivňují názory dalších lidí, a proto také roste význam analýzy sentimentu. Jedná se vlastně o „dolování dat“ (data mining) z textu. Uplatnění analýzy sentimentu může být v různých odvětvích. Analýza sentimentu se začala používat v oblasti informatiky, ale později se ukázalo také její uplatnění například v oblasti sociálních věd, financí nebo managementu.

Tato diplomová práce je zaměřena na analýzu sentimentu v oblasti financí. Nedávné publikace ukázaly, že analýza sentimentu může zpřesnit predikci vývoje na kapitálových trzích. Jedním z klíčových ukazatelů kapitálových trhů je volatilita akciových indexů, respektive jednotlivých akciových titulů. Cílem této práce je za vybrané časové období provést sběr dat z článků týkajících se vybraných firem obchodujících na kapitálových trzích a pokusit se predikovat jejich volatilitu pomocí analýzy sentimentu a dalších finančních ukazatelů.

Práce je rozdělena do šesti hlavních kapitol. První kapitola se věnuje trhu cenných papírů. Tato kapitola se zaměřuje na strukturu trhu, na vlivy ovlivňující ceny akcií, blíže popisuje mimoburzovní systém Nasdaq, na kterém je řada firem vybraných pro diplomovou práci obchodována a v neposlední řadě rozebírá akciové indexy a riziko spjaté s obchodováním.

Následuje kapitola zaměřená na bližší vysvětlení problematiky volatility. Kapitola popisuje modely vhodné pro výpočet volatility a blíže popisuje indexy volatility. Další kapitola je zaměřena na psychologii investování, jelikož toto odvětví úzce souvisí s analýzou sentimentu a obchodováním na kapitálových trzích. Představuje tak teoretický základ pro použití analýzy sentimentu na kapitálových trzích. Čtvrtá kapitola popisuje problematiku analýzy sentimentu, jejíž poznatky budou využity pro analýzu zpráv pro vybrané firmy.

Další kapitola se týká sběru a přípravy dat pro analýzu, kde byly zjišťovány hodnoty sentimentu v článcích volně dostupných na internetu pro 14 největších amerických firem. Hodnoty sentimentu byly získávány pomocí slovníků pozitivních a negativních slov od Loughrana a McDonalda [12]. Následně jsou zde popsány i další ukazatele, které byly použity

k predikci volatility, mezi které patří např. tržní kapitalizace, P/E (tržní cena akcie / zisk na akcii) nebo volume (objem obchodů).

Poslední kapitola je zaměřena na využití výsledků analýzy sentimentu pro predikci volatility vybraných firem. Pro zjištění vlivu sentimentu je v této práci vybrána jedna z nejpoužívanějších statistických metod – vícerozměrná regresní analýza, která je v této kapitole i teoreticky popsána. Před provedením regrese je provedena korelační analýza pro ověření předpokladů regresního modelu. Dále je zde popsán samotný výsledek regresní analýzy, především statistický vliv sentimentu a ostatních faktorů na volatilitu. V závěru kapitoly jsou ověřeny zbývající předpoklady regresního modelu a dosažené výsledky jsou znázorněny graficky.

1 TRH CENNÝCH PAPÍRŮ

Na začátku bude řeč o trhu cenných papírů (CP), jelikož tato diplomová práce se bude zabývat analýzou sentimentu právě na trhu s akciemi amerických společností. Na trhu CP jsou peněžní prostředky alokovány od přebytkových jednotek k jednotkám deficitním prostřednictvím různých druhů CP [13]. Deficitní jednotky emitují různé druhy krátkodobých (v případě peněžního trhu) nebo dlouhodobých (v případě kapitálového trhu) CP a přebytkové jednotky investují své úspory do těchto instrumentů. Efektivnost výše uvedené alokace je zvyšována přímou či nepřímou asistencí investičních prostředníků ve formě emisních, brokerských, dealerských, poradenských, depotních a dalších investičních obchodů. CP kapitálového trhu se v užším slova smyslu označují jako efekty. Ztělesňují pohledávky nebo podíl, nebo zaručují bezpodmínečný nárok na trvalý důchod (státní obligace, obligace soukromých společností, akcie, hypoteční zástavní listy a podílové listy) [10].

1.2 Struktura trhu cenných papírů

CP lze dělit do různých kategorií, kde hlavní dělení je na trh primární a sekundární. Na primárním trhu se obchoduje s nově emitovanými CP, které jsou poprvé uváděny na finanční trh. Rozeznávají se vlastní emise, cizí emise, veřejná emise a soukromá emise [16]. Oproti tomu na trhu sekundárním se obchoduje s již dříve vydanými CP. Dále lze trh CP dělit na veřejný a neveřejný. Na veřejném trhu se obchoduje s CP obvykle za nejvyšší nabídnutou cenu a obchodovat může každý potenciální zájemce, zatímco na neveřejném trhu probíhají pouze tzv. smluvní obchody, při nichž bývají předmětem CP prodány buď pouze jednomu, případně pouze několika kupcům zároveň, a to na základě individuálně dohodnutých podmínek. Jako poslední dělení trhu lze uvést trhy organizované a neorganizované. Zmíněné dělení spadá pod sekundární veřejný trh, kde na neorganizovaném trhu se obchoduje prostřednictvím bank nebo jiných institucionálních či soukromých obchodníků s CP. Jedná se o tzv. „prodej před přepážku“ neboli o „OTC-trhy“. Namísto toho trh organizovaný je více regulován a je reprezentován burzami nebo případně i jinými (mimoburzovními) licencovanými organizátory veřejného trhu. Jejich činnosti spočívá v tom, že agregují předem neomezenou veřejnou nabídku a poptávku, párují vzájemně si odpovídající obchodní příkazy, čímž umožňují vznik spravedlivých tržních cen, které se v těchto případech označují jako „kurzy“, a zároveň zvyšují likviditu trhu [20]. Právě mimoburzovního sekundárního veřejného organizovaného trhu se bude věnovat tato diplomová práce. Jde v podstatě o elektronickou verzi prezenčního systému řízeného cenami. V tomto systému vystupují dva subjekty: jsou to burzovní obchodníci a tvůrci trhu (market makers), kteří plní funkci

burzovních zprostředkovatelů v aktivní roli, což znamená, že mají pro několik jim svěřených emisí ve vymezeném čase povinnost neustále kótovat nákupní a prodejní kurzy, za které jsou ochotni obchodovat na vlastní účet a riziko. Obchodníci mohou na počítači sledovat aktuální kurzy a kótace tvůrců trhu a poté uzavírat obchody buď s tvůrci trhu, nebo mezi sebou za kótace tvůrců trhu. Nespornou výhodou elektronické verze systému řízeného cenami zůstává jeho vysoká likvidita, která je zajišťována nepřetržitými kótacemi tvůrců trhu. Nevýhodou je pak netransparentnost [21]. Systém řízený cenami je používán právě na americkém mimoburzovním trhu Nasdaq, o kterém bude kapitola níže.

1.3 Mimoburzovní trh Nasdaq

Nasdaq¹ je forma mimoburzovního trhu v USA, který nemá podobu obchodování na parketu. Obchodování na organizovaném mimoburzovním trhu je v USA již od počátku 70. let 20. století zajišťováno prostřednictvím elektronického systému Nasdaq, který má podobu obchodního systému řízeného cenou. V důsledku své schopnosti rychle zavádět nové technologie patří Nasdaq mezi nejrychleji rostoucí trhy na světě. V současnosti se na Nasdaqu obchoduje více než 3000 akciových emisí a přes 2000 emisí dluhových instrumentů. V systému Nasdaq působí asi 500 000 obchodníků a více než 500 tvůrců trhu[21]. V systému Nasdaq hrají rozhodující roli market makers (tvůrci trhu), kteří na principu dvoucestné kótace neustále kótují nákupní a prodejní kurzy, za něž jsou ochotni daný titul koupit či prodat. Pro jeden titul existují minimálně dva, ale zpravidla více tvůrců trhu, kteří si navzájem konkurují. Obchodní systém Nasdaqu funguje na několika různých úrovních. Za nejnižší poplatek je možné získat přístup do nejnižší úrovně Nasdaqu, která poskytuje přísun kurzovních informací distribuovaných informačními agenturami Reuters nebo Quotron. Začátkem 90. let minulého století Nasdaq podstatně rozšířil dobu, během níž je možné obchodovat. Zároveň je také zřízeno zpravodajství v reálném čase. V tomto období byl rovněž otevřen nový tržní segment Nasdaq International, který umožnil obchodovat v londýnském čase. V roce 1998 zahájil Nasdaq úzkou spolupráci s americkou burzou AMEX. Na přelomu tisíciletí Nasdaq otevřel několik mezinárodních segmentů v Kanadě, v Japonsku a v Evropě. Funkcí hlavního indikátoru amerického mimoburzovního trhu plní Nasdaq Composite Index [21]. Jelikož dvě třetiny indexu Nasdaq tvoří technologické firmy, investoři jej často používají jako měřítko pro stanovení finančního zdraví technologických akcií. Nevýhodou indexu Nasdaq Composite je skutečnost, že je více spekulativní než ostatní indexy. Jedním z důvodů je, že společnosti kótované v Nasdaq jsou považované za společnosti s vysokým růstovým potenciálem, toto je

¹ National Association of Securities Dealers Automated Quotations

však často doprovázeno vyšším rizikem takové investice. Proto má index Nasdaq větší volatilitu než ostatní indexy [20].

1.4 Fundamentální vlivy na vývoj ceny akcií

Ceny veřejně obchodovaných akcií se na globálním akciovém trhu vytvářejí kontinuálním způsobem na základě střetávání nabídky a poptávky. Tato nabídka a poptávka po akciových instrumentech je generována portfoliovými investory, přímými investory a nepoctivými osobami, které manipulují s tržními cenami, perou špinavé peníze, nebo nadměrně a nevýhodně obchodují na účet zákazníka. Změny v chování účastníků na globálním akciovém trhu jsou podněcovány nejen kolísáním fundamentálních a kurzotvorných faktorů (např. neočekávané globální nebo regionální události, neočekávané změny úrokových sazeb, inflace, alokace kapitálu a zisková překvapení), ale také demografickými, technologickými, politickými, mocenskými a dalšími okolnostmi. Z výše uvedených důvodů je zřejmé, že tržní cena akcie na veřejném trhu může vzniknout pouze střetnutím nabídky a poptávky, přičemž tomuto objektivnímu procesu utvoření tržní ceny předchází subjektivní ohodnocení konkrétního akciového titulu či akciového segmentu ze strany odlišně motivovaných kupujících a prodávajících, a to v kontextu globálním, regionálním či národního investičního prostředí [13]. Cena (kurz) vytvořená hraje vždy významnou informativní roli. Pro investora cena tvoří významné měřítko atraktivity daného CP, zohledňuje možný výnos z investice, ale i část nákladu. Nezastupitelnou informací je cena obchodovaného instrumentu pro analytiky nebo investiční poradce, kteří se pokoušejí odhalit podhodnocené nebo nadhodnocené tituly [21]. V zásadě existují čtyři přístupy pro analyzování vhodnosti nákupu či prodeje akciových instrumentů:

- fundamentální přístup,
- technický přístup,
- psychologický přístup a
- kombinovaný přístup.

Tato práce je nejvíce zaměřena právě na psychologický přístup, kde ceny akcií jsou ovlivněny pomocí sentimentu na kapitálových trzích.

1.5 Akciové indexy

Dále bude stručně vysvětlen pojem index. Veličina, která kvantitativně popisuje určitou sociálně ekonomickou skutečnost, se nazývá ukazatel. V praxi je však zpravidla lepší vědět, zda hodnota ukazatele znamená určitou změnu oproti téže skutečnosti v minulém období či

v jiné územní či organizační jednotce. To znamená, že uživatele nezajímá jenom jedna hodnota daného ukazatele, ale i její relativní, resp. absolutní velikost ve vztahu k hodnotě téhož ukazatele v jiné situaci. Podílem dvou hodnot téhož ukazatele se získá index, rozdílem pak absolutní rozdíl. Obě tyto míry rozdílnosti jsou rovnocenné a nezastupitelné, ale vzájemně se doplňují [3]. Hlavním posláním akciových indexů je stručně informovat o vývoji a výnosnosti určitého akciového trhu. Kromě toho se akciové indexy používají jako tzv. benchmark, tj. jako měřítko průměrného výnosu daného trhu, vůči kterému je možné poměřovat úspěšnost investování správců fondů [6]. Subjekty, které dosáhnou vyšší výnosové míry, než je výnosová míra vypočtená na základě tržního indexu, mohou být označeny jako „nadvýnosové“ oproti benchmarku. Oblastí, kde je údaj o výnosové míře produkované tržním indexem využíván, je kalkulace beta faktoru jako míry systematického rizika CP, popř. portfolia [21]. Charakteristickým rysem všech nejznámějších indexů je pevná báze. Oproti pohyblivé bázi je její výhodou stabilita výběru a větší reprezentativnost celého trhu, protože do báze nevstupují pouze nejlepší a nejatraktivnější společnosti, ale jsou tam i společnosti méně zajímavé. Některé indexy používají geometrický průměr místo častěji používaného aritmetického.

Indexy se rozlišují podle způsobu výpočtu a rozsahu trhu. Index může být [6]:

- vážený – odráží i relativní velikost trhu s danými akciami,
- nevážený – bere v úvahu pouze ceny akcií.

Průměrování může být:

- aritmetické – sečtením cen jednotlivých akcií a tento součet se podělí počtem akcií,
- geometrické – ceny jednotlivých akcií se vynásobí a součin se poté umocní členem $1/n$ (n je počet akcií zahrnutých v indexu).

Indexy založené na prostém průměru

Nejznámějším a nejstarším indexem založeným na prostém aritmetickém průměru je DJIA (Dow Jones Industrial Average), který se počítá od roku 1887. V současné době je DJIA neváženým aritmetickým průměrem ceny 30 velkých amerických průmyslových podniků. Dalším indexem je například Transportation Average, který obsahuje jedenáct nejvýznamnějších železničních společností. Jeho zakladatelem je Charler Dow, který jednoduše sečetl ceny jedenácti akcií a vydělil číslem jedenáct.

Indexy založené na váženém průměru

U indexů založených na váženém průměru cen akcií jsou vahami počty emitovaných akcií zahrnutých do indexu. Například frankfurtský index DAX se počítá jako vážený aritmetický průměr 30 akciových společností. Dalšími indexy jsou například londýnský index FTSE 100 nebo index S&P 500 [6].

1.6 Riziko

Jak již bylo zmíněno, investování na kapitálových trzích s sebou přináší i rizika. Nyní bude pozornost zaměřena na moderní teorii rizika. V reálném světě nemůže žádný investor s naprostou jistotou předpovědět všechny finanční toky, plynoucí z držení aktiva, a tedy jeho výnosovou míru. Nepředvídatelné události, jež jsou denní skutečností v reálném ekonomickém prostředí, se stávají rizikovými faktory, které vedou ke vzniku nejistoty ohledně budoucích příjmů i výdajů. Racionální investor tedy spíše než o dané konstantní výnosové míře bude uvažovat o pravděpodobnostním rozdělení možných výnosových měr. Toto rozdělení vzniká buď objektivně (na základě historických dat), nebo subjektivně (na základě osobních odhadů investora). Většinou vzniká pravděpodobnostní odhad obou postupů dohromady [10]. V této práci se pozornost zaměří na riziko jednoho samostatně drženého aktiva. Druhou možností může být vytvořit si portfolio aktiv, kde je výpočet rizika složitější (viz níže).

1.6.1 Nediverzifikovatelné riziko

Povaha nediverzifikovatelného rizika je systematická, a proto je imunní vůči technikám snižování rizika pomocí diverzifikace. Zdroji systematického rizika jsou změny v politickém, hospodářském a sociologickém klimatu a jejich dopadu na trhy CP. Ceny téměř všech individuálních CP mají tendenci pohybovat se synchronně, a proto ceny téměř všech akcií, obchodovaných na veřejných trzích, jeví významnou vzájemnou pozitivní korelaci. V důsledku toho jsou pohyby indexu, sledujícího pohyby souboru akcií, obchodovaných např. na NYSE², vysvětlením pro téměř 30 % cenových pohybů u všech zhruba 2000 individuálních CP. Tržby, zisky a tedy i ceny akcií pro společnosti se systematickým rizikem silně závisí na celkové společenské úrovni ekonomické aktivity a celkovém stavu trhu s CP [10]. V této práci se bude pracovat právě s tržním nediverzifikovatelným rizikem. Lze jej nazývat též riziko systematické, jak již bylo zmíněno výše [4]. Zdrojem tohoto rizika je

² New York Stock Exchange (zkráceně NYSE), je americká burza se sídlem ve městě New York City, státu New York. Tato burza je největší burza akcií a derivátů na světě

pravidelné střídání býčích a medvědích trendů³ na kapitálových trzích, které mají tendenci zasahovat takřka všechna aktiva stejným způsobem. Střídání trendů je úzce spjato s hospodářským cyklem dané ekonomiky. Empirické studie ukazují, že změny trendu z medvědího na býčí předcházejí téměř každou recesi ekonomiky [10].

1.6.2 Diverzifikovatelné riziko a analýza portfolia

Určitá rizika lze snížit diverzifikací. Zisky podniků závisejí na schopnosti managementu a na celkovém ekonomickém a právním stavu hospodářství. Obecná úroveň ekonomické aktivity se mění podle hospodářského cyklu. Zisky firem jsou v určitém rozsahu na sobě závislé a diverzifikací nelze toho riziko snížit. Toto riziko jde do určité míry omezit investováním na zahraniční trhy [6]. Jednoduchá neboli prostá diverzifikace je založena na předpokladu, že portfolio, složené z 200 akcií, má desetkrát nižší riziko než portfolio, složené z 20 akcií. Jistým způsobem se takto riziko snížit dá [10]. Velikost rizika portfolia měřena jeho rozptylem či směrodatnou odchylkou je závislá nejen na výši rizika jednotlivých složek portfolia, ale také na míře jejich statistické závislosti. Přímá statistická závislost neboli pozitivní korelovanost složek přispívá ke zvýšení rizika portfolia. K vyváženému portfoliu lze dospět takovým výběrem projektů, který zahrnuje především statisticky nezávislé projekty [4].

1.6.3 Celkové riziko

Celkové riziko je složené ze systematické a nesystematické složky. Systematické riziko tvoří asi čtvrtinu celkového rizika akcie. Zbývající nesystematická složka rizika může být neutralizována rozložením investic do většího množství akcií, nicméně v okamžiku, kdy je celkové riziko sníženo na úroveň systematického, další zvyšování počtu akcií v portfoliu ke snížení rizika nevede [10].

Celkové riziko, které je spojeno s držením kapitálového aktiva, je dáno celkovou variabilitou jeho očekávaných výnosů. Míru rizika spojeného s aktivem lze uvažovat jako míru variability výnosů, kterou je rozptyl a jeho druhá odmocnina, která se nazývá směrodatnou odchylkou [10]:

$$\sigma = \sqrt{\text{var}(r)}. \quad (1)$$

³ Každý trh, který roste (cena komodity jde nahoru), se nazývá býčí trh. Naopak trh, který klesá (cena komodity jde dolů), se nazývá medvědí trh.

1.6.3.1 Určování rizikovosti prostřednictvím statistických nástrojů

Výpočtů rizikovosti finančních investic prostřednictvím statistických nástrojů se používá především tehdy, je-li k dispozici dostatečné množství vhodných statistických údajů. Těmi bývají nejčastěji různé dlouhodobé ekonomické ukazatele, jež je možno uspořádat do dostatečně dlouhých průkazných časových řad. Je pochopitelné, že statistika disponuje celou řadou různých statistických charakteristik, z nichž se v praxi při určování tzv. „absolutní výše rizika“ nejčastěji používá „rozptylu“ a „směrodatné odchylky“. Právě tyto ukazatele vyhovují nejlépe poznání, že nestačí pouze určit „očekávanou výnosnost“ zvažované investice, ale zároveň i to, jak se od ní mohou „skutečně dosažené hodnoty výnosnosti“ odchýlovat. Směrodatná odchylka je statistickým odhadem pravděpodobného odchýlení se „skutečné výnosnosti“ předmětné finanční investice od její „výnosnosti očekávané“. Je druhou odmocninou rozptylu neboli kvadratického váženého průměru odchylek od očekávané míry výnosu, v němž jsou vahami pravděpodobnosti těchto odchylek, jejichž součet se musí rovnat jedné. Proto obecně platí, že čím bude hodnota směrodatné odchylky větší, tím větší bude i rizikovost předmětné investice [4].

Vzorec pro výpočet směrodatné odchylky od očekávané míry výnosu [20]:

$$\sigma = \sqrt{\frac{\sum_1^n p_i \cdot (r_i - E(r_i))^2}{\sum_1^n p_i}}, \quad (2)$$

kde: σ je směrodatná odchylka.

r_i je hodnota výnosnosti,

$E(r_i)$ je očekávaná výnosnost,

p_i jsou pravděpodobnosti jednotlivých odchylek,

n je počet aktiv,

i je zvolené aktivum.

2 VOLATILITA

Volatilita je důležitým ukazatelem na finančních trzích. Mnoho lidí mylně volatilitu zaměňuje s rizikem. Každý investor má jiný práh snesitelnosti rizika. Správná předpověď volatility při oceňování aktiv v průběhu jejich držení je dobrým začátkem pro posuzování investorského rizika [18]. Predikci volatility lze považovat za významnou úlohu finančního modelování. Je tomu tak proto, neboť volatilita je základním parametrem při řízení finančních rizik, promítá se při oceňování opcí, optimalizaci portfolia, uplatňování metody Value at Risk, hedgingu apod. [23]. K ocenění opce je třeba znát volatilitu podkladového aktiva od začátku až do vypršení lhůty opce. V dnešní době existují speciální druhy derivátů, při jejichž koupi je ve smlouvě změřena a definována volatilita, a proto je nezbytným pomocníkem při určení ceny derivátů [18]. Pro stanovení běžné volatility lze využít historického přístupu. Vzhledem k tomu, že často nebývá splněn předpoklad homoskedasticity⁴, je nutné aplikovat adaptační metody založené na stanovení podmíněného rozptylu. Dobré zkušenosti jsou s modelem GARCH⁵ a modelem EWMA [23]. Volatilita se využívá také pro vyjádření vztahu nejistoty na finančních trzích a důvěry veřejnosti. Lze ji nazvat jako barometr pro měření zranitelnosti finančních trhů v závislosti na ekonomice státu [18]. Přestože volatilita není přímo pozorovatelná, má určité charakteristiky, které jsou obvyklé, když se právě sleduje výnosnost nejrůznějších finančních aktiv. Tyto charakteristiky jsou následující [8]:

- Shlukování volatility: volatilita může být v některých obdobích vysoká a v jiných nízká.
- Pákový efekt: volatilita reaguje odlišně na cenový vzestup a na cenový pokles.
- Kontinuita: volatilita se vyvíjí spíše spojitě bez nějakých výrazných skoků.
- Omezenost: volatilita nesměřuje k extrémně vysokým (neomezeným) hodnotám, ale její průběh bývá spíše stacionární v určitém rozmezí.

Nejjednodušeji lze volatilitu vyjádřit jako směrodatnou odchylku ceny aktiva σ , nebo jako rozptyl σ^2 , odvozené ze vzorce [18]:

$$\sigma^2 = \frac{1}{N-1} \sum_{t=1}^N (r_t - \bar{r})^2, \quad (3)$$

⁴ Homoskedastičnost ve statistice znamená, že rozptyl je nezávislý na parametru (je homogenní).

⁵ GARCH model odstraňuje nedostatky modelu ARCH, ze kterého je odvozen. Model ARCH je zaměřen na modelování volatility. Jedná se patrně o nejvýkonnější finanční nástroj pro modelování časových řad. [8]

kde: t je čas,

N je zvolené časové období.

Důležitým indikátorem, který vychází z volatility je tzv. koeficient beta. Ten nabývá hodnot od 0 do nekonečna a značí relativní volatilitu daného aktiva či portfolia vzhledem k nějaké základně (například akcie a akciový index). Značí tedy, jak moc spolu koreluje volatilita indexu a konkrétní akcie [5]:

Beta = 0 až 1 - pokud je beta v rozmezí od 0 do 1 (příčemž nezahrnuje ani jednu z krajních hodnot), je volatilita zkoumané akcie menší než indexu;

Beta = 1 - v tomto případě jsou volatility naprosto shodné;

Beta > 1 - zde je akcie volatilnější než index samotný.

2.1 Model oceňování kapitálových aktiv (CAPM)

Koeficient beta je neodmyslitelně spojen s modelem CAPM. Tento model předpokládá nulové transakční náklady na diverzifikaci, nekonečnou dělitelnost aktiv, dokonalou informační symetrii, nemožnost najít nesprávně ohodnocená aktiva a existenci dokonale bezrizikového aktiva (v praxi jsou za něj považovány státní dluhopisy). Logickým důsledkem současné platnosti těchto předpokladů je, že všichni investoři budou držet kombinaci tržního portfolia (dokonale diverzifikovaného mezi všechny finanční i reálná aktiva v ekonomice, z důvodu odstranění jedinečného rizika) a bezrizikového aktiva. Jediné riziko v modelu CAPM je z těchto důvodů systematické riziko [5]. Vztah mezi očekávanou výnosovou mírou a systematickým rizikem vyjadřuje přímka trhu cenných papírů (security line – SML) viz obrázek č. 1, kterou lze matematicky zapsat následujícím způsobem[20]:

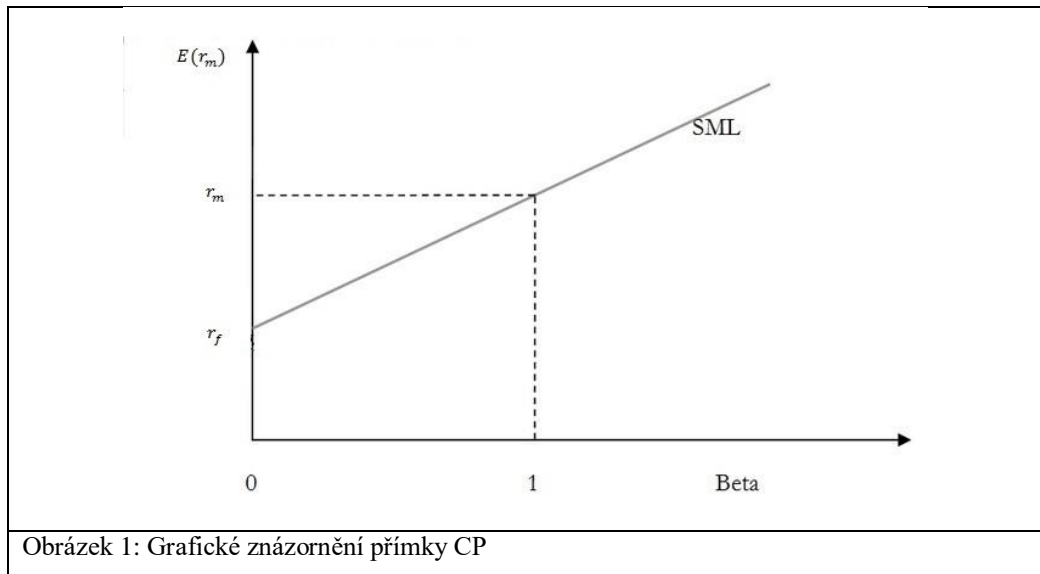
$$E(r_i) = r_f + \beta_i [E(r_m) - r_f], \quad (4)$$

kde: $E(r_i)$ je očekávaná výnosová míra aktiva i ,

r_f je bezriziková výnosová míra ze státních pokladničních poukázek,

$E(r_m)$ je očekávaná výnosová míra z tržního portfolia,

β_i je beta faktor, který vyjadřuje citlivost i -té investice na změnu výnosové míry z tržního portfolia.



Zdroj:[13]

Beta faktor i -tého aktiva vyjadřuje citlivost tohoto aktiva na změnu výnosové míry tržního portfolia, což lze zapsat jako [13]:

$$\beta_i = \frac{cov_{im}}{\sigma_m^2}, \quad (5)$$

kde β_i je beta faktor,

cov_{im} je kovariance mezi výnosovou mírou i -té akcie a výnosovou mírou z tržního portfolia,

σ_m^2 je rozptyl výnosové míry z tržního portfolia.

2.2 Druhy volatility:

Výše byla zmíněna běžná volatility, která se měří, jak již bylo řečeno, pomocí historických dat. Lze definovat i další typy volatility [22]:

- **historická volatility** - volatility finančního instrumentu, která odkazuje na minulá data, včetně posledního známého údaje (oproti běžné volatilítě, která má poslední údaj aktuální);
- **budoucí volatility** - volatility finančního instrumentu, která se začíná počítat dnešním datem a končí v budoucnu (typicky končí u data expirace opce);
- **korelovaná volatility (Beta)** - porovnává volatility daného aktiva či portfolia vůči benchmarku (například akciový index);

- **implikovaná volatilita** - volatilita finančního derivátu (v drtivé většině opce), která je odvozena od volatility podkladového aktiva; bývá jedním ze vstupů do oceňovacích modelů opcí (např. Black-Scholes).

2.3 Riziko volatility

Riziko volatility nebo také tržní riziko se projevuje vždy, když je v portfoliu libovolné množství cenných papírů (akcií, dluhopisů) stejného druhu, odvětví, regionů apod. Tehdy je jedno, zda se jedná o akcie jedné firmy, nebo více firem. Riziko volatility působí na všechny podobně. Do této kategorie se řadí i rizika výkyvů ekonomiky, přírodní katastrofy, politické převrasy atd. Riziko volatility lze tedy diverzifikovat pouze investováním do různých investičních příležitostí (např. do akcií firem různých odvětví, do dluhopisů (různých států), do finančních derivátů, komodit, nemovitostí apod.) [18].

2.4 EWMA model

Pragmatickým rozšířením předchozího přístupu jsou modely EWMA⁶. Nejpoužívanější EWMA-model volatility představuje analogii jednoduchého exponenciálního vyrovnávání pro volatilitu. Na rozdíl od výpočtu historické volatility se zde při průměrování váží tím způsobem, že váhy klesají geometricky do minulosti. To má ve srovnání s historickou volatilitou řadu praktických předností:

- V praxi bývá volatilita skutečně více ovlivněna aktuálními pozorováními, které jsou v EWMA-modelu zvýrazněny většími vahami, než pozorováními vzatými z hlubší minulosti, které mají v EWMA-modelu nižší váhy.
- V EWMA-modelu se samovolně redukuje problém odlehlého pozorování s abnormální velikostí.

Volatilita v modelu EWMA se odhaduje jako [8]:

$$\hat{\sigma}_t^2 = (1 - \lambda) \sum_{j=0}^{\infty} \lambda^j (y_{t-j} - \bar{y})^2 = (1 - \lambda)(y_t - \bar{y})^2 + \lambda \hat{\sigma}_{t-1}^2, \quad (6)$$

kde odhadnutá volatilita $\hat{\sigma}_t^2$ je zároveň předpovědí (nepříliš vzdálené) budoucí volatility z času t , \bar{y} je průměrná úroveň dané řady a λ ($0 < \lambda < 1$) je předem zvolená diskontní konstanta.

⁶ Exponentially weighted moving averages

2.5 Indexy volatility

Jde o poněkud zvláštní, avšak postupně stále významnější indexy, označované také jako „indexy strachu“ či „indexy paniky“. Jsou měřítkem očekávané volatility pro kupní a prodejní opce na nejvýznamnější podkladové akciové indexy, jako např. na akciový index S&P 500, Nasdaq Composite a další významné obchodovatelné akciové indexy. Charakteristickou vlastností těchto indexů je to, že mají inverzní vztah k akciovým indexům. To znamená, že pokud indexy volatility rostou, tak akciové indexy klesají a naopak. Je to způsobeno tím, že zvyšování hodnoty indexu volatility znamená, že roste nervozita investorů, a pokud dosáhne extrémních hodnot, znamená to, že nastala panika. Indexů volatility existuje několik a záleží na jejich konstrukcích a mnoha dalších souvisejících faktorech, jaké dávají výsledky a doporučení investorům [18].

Volatility S&P 500 index (VIX)

Jedná se o nejstarší a nejpoužívanější index volatility na světě. Je počítán pro kupní a prodejní opce obchodované na CBOE a prezentuje očekávanou třicetidenní volatilitu této burzy obchodující opce na akciový index S&P 500. VIX bývá charakterizován jako barometr sentimentu investorů a volatility trhu. Jako extrémní hodnoty jsou u něj považovány hodnoty nad 30, resp. 35 body, což značí, že panika investorů dosáhla maxima. Za minimální hodnoty se považuje hladina kolem 10 bodů, kdy na trzích vládne důvěra v budoucí pozitivní vývoj [20].

Další indexy volatility

Jako další indexy volatility, které CBOE kalkuluje, lze uvést následující [20]:

VXN, jenž souvisí s obchodováním opcí na index Nasdaq 100,

VXD, který se kalkuluje na volatilitu opcí na index DJIA,

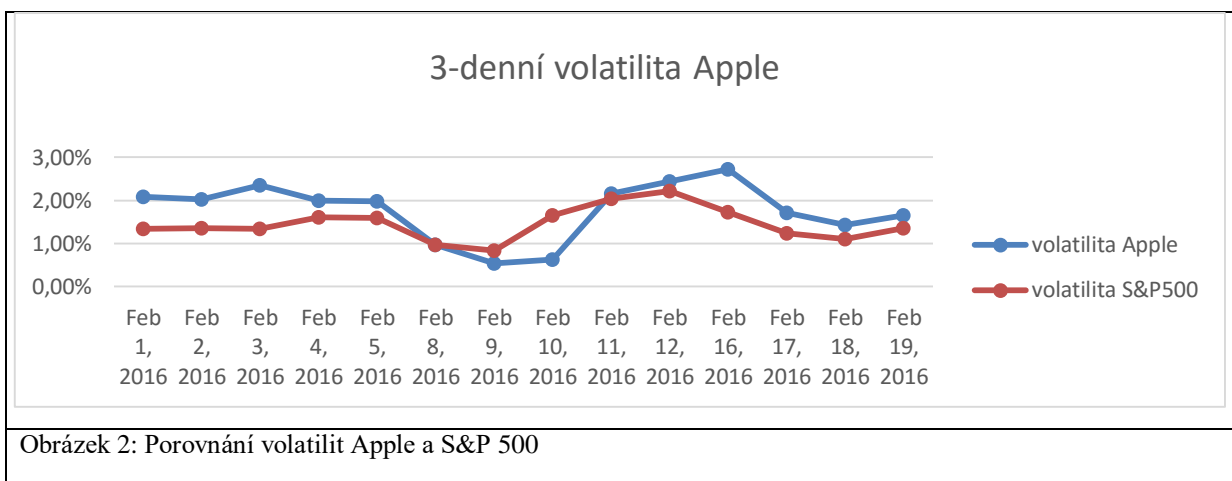
VXO, jenž se odvíjí od opcí na index S&P 100,

aj.

Pro ukázkou lze uvést vypočtené 3-denní volatility všech firem vybraných pro analýzu sentimentu a porovnat je s indexem S&P 500, jenž byl zmíněn v předešlé kapitole. Následně bylo vybráno několik nejzajímavějších firem (zbývající jsou v příloze A). Pokud byla

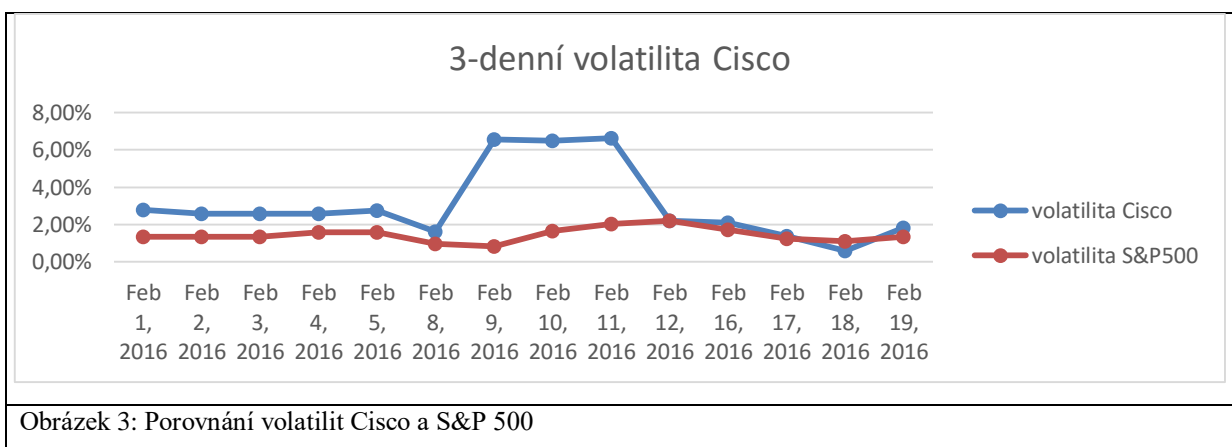
volatilita vyšší, než volatilita indexu, znamená to, že akcie firmy je volatilnější než celkový tržní index S&P 500 a naopak.

Graf porovnání firmy Apple a indexu S&P 500 je zajímavý, jelikož volatilita firmy Apple jako jedna z mála celkem věrně kopíruje trend indexu S&P 500. Celkově se ale dá říci, že Apple je volatilnější než celkový tržní index S&P 500 (obrázek č. 2).



Obrázek 2: Porovnání volatilit Apple a S&P 500

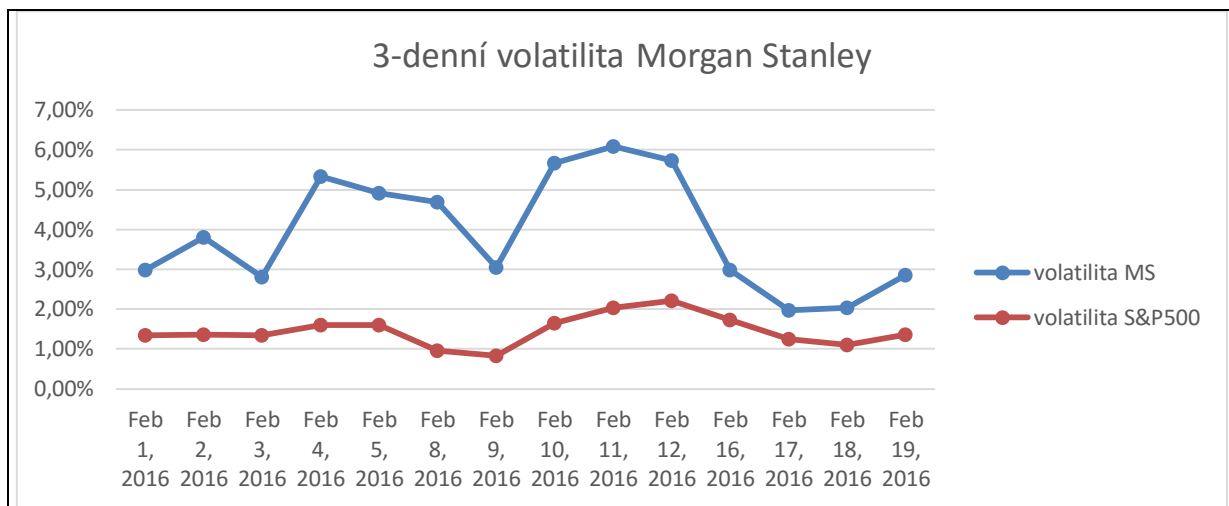
Zdroj: Vlastní zpracování



Obrázek 3: Porovnání volatilit Cisco a S&P 500

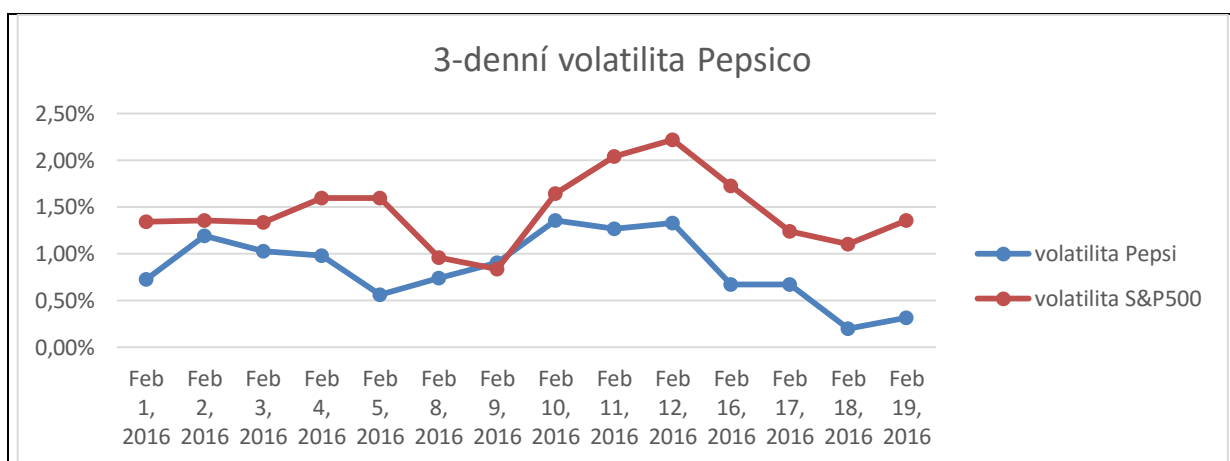
Zdroj: Vlastní zpracování

I firma Cisco kopíruje v určitém období index S&P 500, zajímavý je však náhlý nárůst volatility v období od 8. do 11. února (obrázek č. 3). U volatility firmy Morgan Stanley je očividné, že je v průběhu sledovaného období stále mnohem volatilnější než index S&P 500. Dále můžeme vidět i velkou nestabilitu v trendu volatility firmy Morgan Stanley (obrázek č. 4).



Obrázek 4: Porovnání volatilit Morgan Stanley a S&P 500

Zdroj: Vlastní zpracování



Obrázek 5: Porovnání volatilit Pepsi a S&P 500

Zdroj: Vlastní zpracování

Naopak u firmy Pepsi je vidět, že volatilita indexu S&P 500 je až na jeden bod vždy vyšší než volatilita firmy Pepsi. Jejich trend je jinak velmi podobný (obrázek č. 5).

3 PSYCHOLOGIE INVESTOVÁNÍ

Jelikož je tato diplomová práce úzce spjata s psychologií investora, bude tomuto tématu věnována celou kapitola. Na investora působí celá řada vnějších vlivů, které ovlivňují jeho obchodní jednání. Tato práce se zaměřuje na mediální články, které investoři čtou a na základě subjektivního vnímání těchto informací dále jednají. Důležité je vyhodnocení obsahu článku, zda na investora celkově článek působil spíše pozitivně nebo negativně. Pro odhalení sentimentu v textu bude použita analýza sentimentu, která tzv. doluje data z textu. Analýze sentimentu se bude věnovat další část práce, a proto budou nyní představeny i jiné vnější i vnitřní psychologické vlivy, které investora ovlivňují.

3.1 Interní chyby investora

Základní emoce, které ovládají myšlení investora a jeho postoj k riziku, jsou naděje a strach. Strach nutí investory zaměřit se na události, které jsou obzvlášť negativní, zatímco naděje je svádí soustředit se na události, které jsou pozitivní. Dalším důvodem, který mění ceny akcií je chybovost investorů i těch profesionálních. Chybovat je lidské, ale investoři své chyby opakují. Nyní budou představeny největší a nejvíce opakované chyby investorů v jejich myšlení [2].

Přílišná sebedůvěra

Psychologové odhalili, že většina lidí trpí velkou sebedůvěrou a v důsledku toho podceňují riziko a přeceňují své znalosti a dovednosti. Přílišná sebedůvěra vede ke špatné interpretaci získaných informací a dále k přeceňování vlastních schopností informace analyzovat. Tento typ investora má tendenci příliš často obchodovat se svým portfoliem. Výše obrátky portfolia bývá spíše negativně korelována s výnosy portfolia.

Dispoziční efekt a odpor ke ztrátám

Lidé jsou disponováni k tomu, že vyhledávají věci a činy, které jim přinášejí pocit pýchy a mají odpor k věcem a činům, které jim způsobují pocit lítosti. U investování se tato vlastnost projevuje tak, že investor má problém prodat investici, u které je ve ztrátě. Domnívá se, že by tím jen potvrdit dřívější špatné rozhodnutí a má pocit, že nerealizovaná ztráta není ztráta. To samozřejmě vede k situaci, že investor prodává spíše investice, které jsou ziskové. Paradoxně se tak zbavují příliš brzy výnosných investic a příliš dlouho lpí na investicích ztrátových.

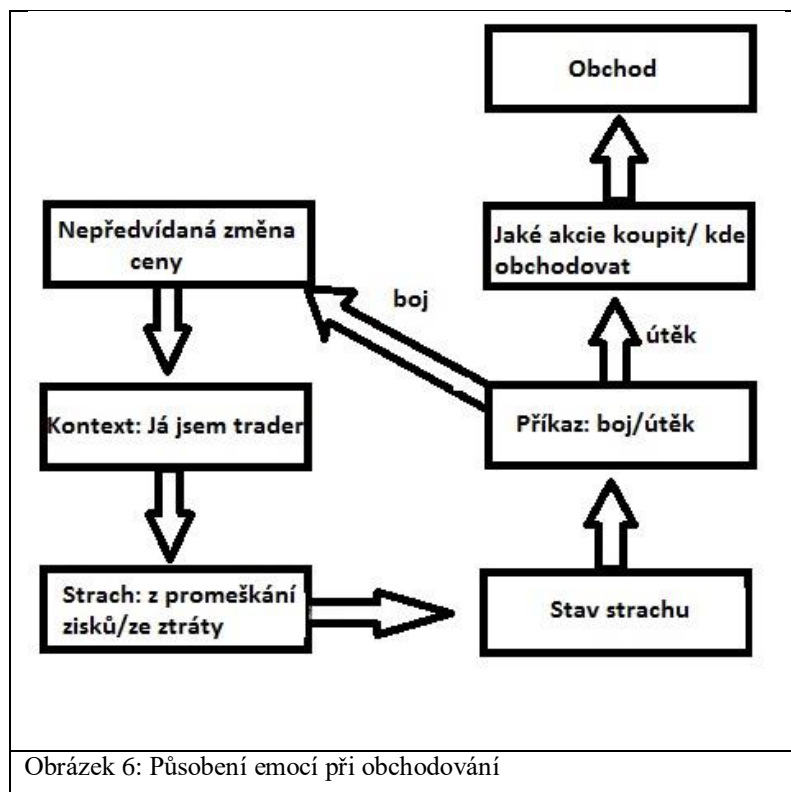
Mentální účtování

Společnosti používají k evidenci svých výsledků skutečné účetnictví. Lidé, jedinci, používají tzv. mentální účtování. Jednoduše řečeno jde o nálepkování peněz. Lidé své peníze rozdělují zvláště např. peníze na byt, peníze na dovolenou atd. Díky tomu se pak nedokáží dívat na svůj majetek jako na celek a dělají iracionální rozhodnutí. Investor má tedy tendenci dávat své investice zvláště do samostatných přihrádek a nedoceňuje tak jejich vzájemný dopad.

Iluze platnosti a naivní extrapolace

Lidé často věří věcem a vztahům, které nemají žádnou platnost. Dělají to proto, že vyhledávají evidenci, která potvrzuje jejich víru a domněnky a ignorují evidenci, která jejich víru a domněnky popírá. Tím podléhají iluzi platnosti, která jde často ruku v ruce s přílišnou sebedůvěrou. S touto vlastností je úzce spojen konzervatismus. Jde o to, že lidé jen pomalu mění své názory, zvyky a to, v co věří. Investoři, kteří touto vlastností trpí, pak odmítají přijímat nová fakta. S iluzí platnosti úzce souvisí také naivní extrapolace. Investoři mají tendenci naivně prodlužovat trendy z poslední doby. Když trh v poslední době roste, mají tendenci domnívat se, že poroste i nadále. Největší vliv má čerstvá historie. Poslední dva měsíce vývoje trhu mají nejsilnější dopad na uvažování o vývoji budoucím. Tento postup často vede k předražení rostoucích akcií a k podhodnocení klesajících akcií a soustředění se na spekulativní odhady vývoje cen více než na vyhledávání akcií hluboko pod svou vnitřní hodnotou.

Z výše uvedeného plyne, že k neadekvátním reakcím dochází zejména v situacích vyvolávajících úzkost a strach. Naše základní reakce na úzkost a strach byly naprogramovány do našich duší ve velmi raném věku a většina z nás si je ani neuvědomuje. Úzkost je velmi snadno vyvolána v okamžiku kontaktu s obchodními a investičními pozicemi na finančních trzích [17]. Další vrozené chyby investorů pramení z přílišné sebedůvěry, ze špatné manipulace se svými investicemi a například z odporu ke ztrátám. Vlivů je samozřejmě mnohem více a každý investor je jiný, proto nemůžeme říci, že všichni reagují stejně podle stejných pravidel. Na obrázku č. 6 je vidět proces působení emocí při obchodování.



Zdroj: [17]

3.2. Externí chyby

Nyní budou představeny externí chyby, které ovlivňují chování investora jako například informace od makléřů, profesionálních analytiků, informace z médií a z internetu. Rozlišovat mezi takovým množstvím informací zda se jedná o relevantní, irelevantní, laické či profesionální informace je velmi těžké a je k tomu zapotřebí odvaha. Pozornost bude zaměřena na dvě externí chyby, které nejvíce souvisí s touto prací.

Vliv médií

Role médií v posledních desetiletích výrazně roste s tím, jak se rozrostl jejich dosah a rychlost přenosu zpráv. Tato skutečnost s sebou přináší řadu pozitiv, ale také řadu negativ. Většina médií má komerční charakter a k tomu také podřizují své chování. Problémem médií je i fakt, že se snaží z každé zprávy udělat velkou věc. Vede investory přemýšlet v krátkých investičních horizontech a k příliš častým transakcím. Dále se média snaží vyvolávat spíše neopodstatněný optimismus než realismus. Například stanice CNBS přináší mnoho velmi kvalitních informací a je asi nejlepším masovým médiem týkajících se investic. Bohužel konflikt zájmů pramenící z komerčního zaměření stanice s sebou přináší i škodlivé prvky.

Právě stanice CNBS je jedním z mnoha mediálních zdrojů, které jsou zahrnuty do analýzy sentimentu v této práci.

Vliv internetu

Velké množství informací a rychlost s jakou informace na internet přicházejí, může investora pomýlit dvojnásobem. Za prvé, bude čelit nebezpečí, že podlehne klamu přílišné sebedůvěry. Za druhé bude v pokušení reagovat na rychlost internetu rychlostí svých rozhodnutí. Práce s internetem vyžaduje velkou sebekontrolu. Dalším nebezpečím je anonymita dostupných informací [2]. Kromě interních a externích vlivů záleží i na typu investora, jelikož každý investor reaguje jiným způsobem na vzniklou situaci na trhu.

3.3 Typy investorů

Konfliktní investor

Tento typ investora je velmi úzkostlivý ohledně svých investic. I když zrovna neinvestuje, stále je myšlenkami u svých obchodů. Přemýšlí, zda neměl něco udělat jinak. V jednu chvíli má ze svých obchodů skvělý pocit a v další minutě může otočit a myslet si, že možná udělal chybu. Je to velmi nevyvážený a rozporuplný typ investora.

Depresivní investor

Depresivní investor se cítí velmi úzkostlivě a špatně i v případě, že mu obchod vyšel. Pokud tento investor pokazí obchod, jen si vlastně potvrzuje své neuvážené jednání. V opačném případě si vyčítá, že do úspěšného obchodu nevložil více peněz. Jen zřídka si depresivní investor vyplatí vůbec nějaké peníze.

Maskovaný investor

Maskovaný investor chce být nejlepším ze všech. Vítězství v obchodu je jediné, čeho chce dosáhnout. Často chce maskovaný investor být někým jiným, lepším než je on sám. Investiční chování je úzce spjaté s aktuální náladou investora.

Úzkostlivý investor

Úzkostlivý investor je velmi zapleten do svého obchodování a zaměřuje se na detaily. Každý krok jeho obchodování zvažuje velmi hluboce. Jedná se o typ investora, který má všechno dokonale srovnané a má ve všem naprostý pořádek [1].

4 ANALÝZA SENTIMENTU

Analýza sentimentu (jinak řečeno také „dolování názorů“) je obor, který zkoumá a analyzuje názory, sentiment, hodnocení, postoje a emoce lidí vůči subjektům jako jsou produkty, služby, organizace aj. V průmyslu se více používá jako název analýza sentimentu, avšak na akademické půdě jsou názvy analýza sentimentu a dolování názorů využívány stejně. V podstatě oba názvy reprezentují stejný vědní obor. Analýza sentimentu se zaměřuje na negativní či pozitivní celkový sentiment lidí. V dnešní době se čím dál tím více lze setkat s velkou škálou různých názorů na internetu [11].

Názor je výraz, který se skládá ze dvou částí:

- z cíle (který můžeme nazvat téma),
- a ze sentimentu, který souvisí s tématem.

Například ve větě „Miluji tuto firmu.“ je sloveso miluji vyjádřením sentimentu k tématu nějaké konkrétní firmy [15]. Právě díky velkému množství informací, které lze na internetu vyhledat, se analýza sentimentu stává více potřebnější než dříve. Je možné ji běžně provádět také díky rostoucímu počtu dostupných dat. Právě analýza sentimentu je nyní centrem sociálních a mediálních výzkumů. Analýza sentimentu má hluboký dopad například na odvětví managementu, politickou sféru či ekonomiku. Výzkumy pomocí analýzy sentimentu začaly především na začátku roku 2000. Je tedy patrné, že se jedná o nový obor [11]. Právě nepřehledné množství informací na internetu začíná být problémem pro zákazníka, který si například chce koupit konkrétní produkt a ve velkém množství informací a diskuzí o tomto produktu se může ztrácet. Je pro něho těžké vyselektovat relevantní informace a udělat správné rozhodnutí o svém výběru [11]. Analýza sentimentu je významná pro pohyby cen akcií a obratu. V literatuře [7] jsou důkazy o bezprostředních účincích sentimentu na kapitálové trhy. V zásadě negativní sentiment tlačí tržní ceny akcií dolů a naopak [18]. Bylo zjištěno, že negativní sentiment v médiích ovlivňuje ceny spíše krátkodobě než trvale. Z výše uvedeného není pochyb, že sentiment v textech má predikční schopnosti v pohybech na trhu. Jako příklad věty, která v sobě skrývá sentiment, si lze uvést následující příklady [19]:

“Miluju svůj nový Ipad”

“Tento film je ten nejlepší, který jsem kdy viděl”

“Nejhorší bageta ve městě, jakou jsem kdy ochutnala“

Určitý pojem si každý člověk zařadí do svého sémantického prostoru, z něhož se odvodila jednoznačná metrika pro určování významu slov. Ty je potřebné posuzovat ze tří základních pohledů. Jde o tyto tři emoční dimenze psaného textu [15]:

- Hodnocení (*posouzení pojmu pomocí bipolárních adjektiv; pozitivní – negativní*)
- Potence (*účinek výpovědi je silný – slabý*)
 - distance (*vztah autora k tématu*)
 - specifičnost (*forma formulace: jasná, vágní*)
 - určitost (*autor si je jistý nebo je na pochybách?*)
- Intenzita (*emotivnost výpovědi*)

Tyto tři kategorie a jejich podkategorie lze rovněž s výhodou použít pro automatickou detekci a klasifikaci pocitů a určení sentimentu konkrétního textu – např. počítačem [12]. Analýza sentimentu se soustředí na posouzení postojů a názorů na určité téma nebo objekt pomocí techniky strojového učení. Dále se analýza sentimentu vysvětluje dle definice dvojího rozdělení na funkční a provozní. Funkční rozdělení se na analýzu sentimentu dívá z praktického hlediska využití.

Při tvorbě analýzy sentimentu je 5 hlavních faktorů, které se musí v textu identifikovat [15]:

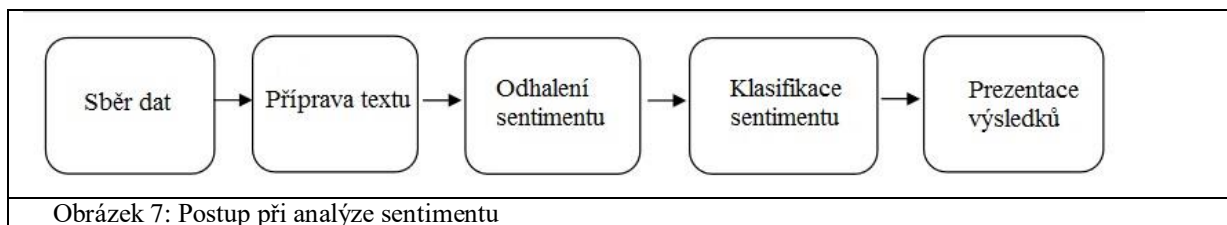
- **Téma** (Co je hlavním předmětem diskuze?)
- **Aspekty, atributy** (O čem konkrétně se v tématu bude hovořit?)
- **Sentiment** (Jaký sentiment se skrývá v názorech?)
- **Majitel názoru** (Kdo názor vyjadřuje? Je více majitelů názorů v textu? A pokud ano, shodují se nebo ne?)
- **Čas** (Kdy byl text pořízen? Nejedná se o zastaralou informaci?).

Literatura spojená s marketingem uvádí dva typy výzkumných metod pro analýzu sentimentu: kvantitativní a kvalitativní. Kvantitativní metody se využívají v případě, kdy se ověřuje nějaká hypotéza. U kvantitativní metody se sbírá velké množství vzorků dat od respondentů. Kvantitativní přístup je bohužel považován za velmi rigidní, jelikož nebere v potaz subjektivitu lidských interakcí. Na druhou stranu je tento přístup lepší pro posouzení celkové reality ve společenském prostředí. Kvantitativní přístup dokáže rozpoznat postoje, víru nebo například emoce vůči určitému objektu. Kvalitativní přístup je považován za velmi

subjektivní a nevědecký, jelikož postrádá strukturu. Nedávný vývoj metodologie naznačil, že by se oba přístupy měly sloučit v jeden komplexní [19].

4.1 Postup při analýze sentimentu

Následující obrázek č. 7 znázorňuje postup při analýze sentimentu.



Obrázek 7: Postup při analýze sentimentu

Zdroj:[19]

1. Sběr dat

Dat pro analýzu sentimentu je na internetu nepřehledné množství. Stránky, které tyto data poskytují, jsou například různé blogy, internetové diskuze, recenze na produkty a služby, Facebook nebo Twitter. Problémem těchto dat je velká neuspořádanost, jelikož každý respondent používá jiný druh vyjadřování, a proto jsou někdy stejné pocity vyjadřovány různými slovy. Díky této skutečnosti není možné provádět analýzu sentimentu manuálně. Právě inovativní textová analytika a přirozené zpracování jazyka se používá k extrakci a klasifikaci těchto dat. Jakmile jsou všechna data extrahována, jsou připravena k analýze [19].

2. Příprava textu

Tato část zahrnuje očištění sesbíraných dat, především pak je cílem identifikovat a odstranit z dat slova neobsahující text. Dále se odstraní ta část textu, která by mohla prozradit totožnost respondenta. V zásadě se odstraní ta slova, která nejsou relevantní pro cíl zkoumání.

3. Odhalení sentimentu

Pomocí výpočetní techniky se analyzuje textová datová sada a automaticky se posuzují názory. V datovém souboru jsou ponechány jen ty věty, které obsahují subjektivně zabarvená slova. Věty, které obsahují fakta a objektivní názory, jsou vyřazeny. Analýza se provádí na různých úrovních. Lze analyzovat termín, jednotlivá slova, celé věty nebo celý dokument s běžně používanými technikami jako jsou [19]:

- Unigram – jde o klasický přístup, kdy je každý prvek reprezentován jako příznakový vektor založený na frekvenci jediného slova. Často je unigram popisován jako „škála slov“.
- *N*-grams – v tomto přístupu je každý prvek reprezentován frekvencí více slov nebo spojení najednou. *N*-grams proto objeví mnohem více kontextu mezi slovy než předchozí přístup.
- Lemmas – přístup lemmas hodnotí slova synonymem spíše než doslovným překladem. Například slova lepší a skvělý vyhodnotí jako dobrý u obou verzí. Díky přístupu lemmas se klasifikace stává jednodušší, jelikož pomáhá se zobecněním slov.
- Negation (negování) – jedná se o rozšíření *N*-gramu, kdy se dvě věty, které se navzájem negují, roztrídí do opačných skupin.
- Opinion words (list názorů) – jde o slova, která vyjadřují lidské pocity a názory (slovesa, přídavná jména, podstatná jména atd.) Tato slova jsou poté zařazena do specifického vektoru. Právě tato zbarvená slova přispívají k lepšímu hledání sentimentu v dokumentech a textech.

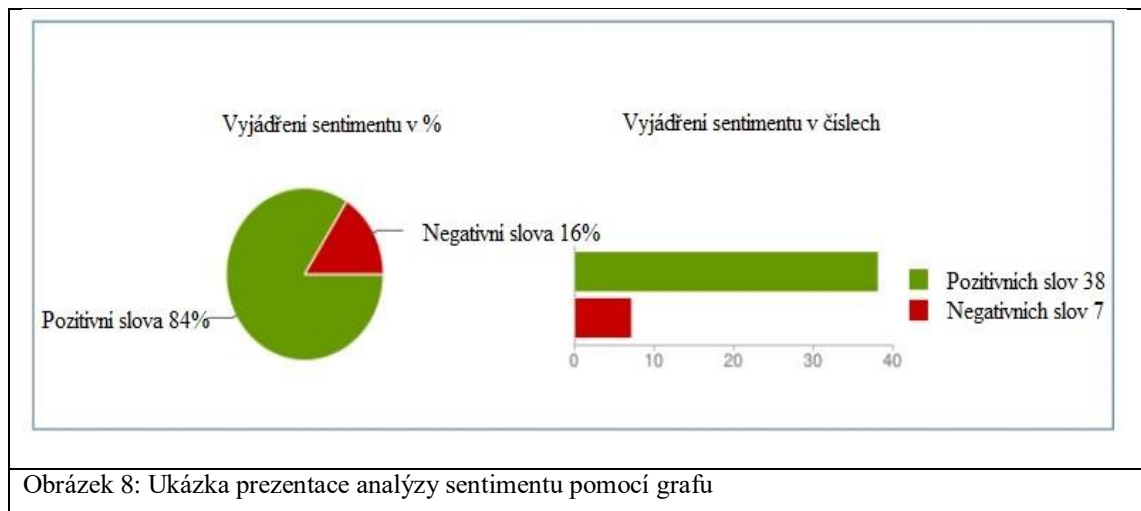
4. Klasifikace sentimentu

V této fázi se musí zjištěný sentiment v textu klasifikovat do vytvořené kategorie. Většinou se jedná o velmi protichůdné skupiny typu dobré, špatné, líbí se, nelíbí se atd. Nicméně klasifikace může být pojata i podobným stylem jako je v hotelech nebo restauracích pomocí hvězd. Obecným cílem je pomocí vložených vstupů či atributů (nezávislá proměnná x) predikovat závislou proměnnou y . Nezákladnější formou klasifikace je „škála slov“ (bag of words). Princip funguje tak, že pro celý dokument se určí počet kladných a záporných slov. Bohužel se tato metoda nehodí pro velký objem dat. Výhodou je však její jednoduchost. Další přístupy zahrnují využití slovníků. Slovníky jsou jakýmsi „mostem“ mezi jazykem a správným vyjádřením a pochopením tohoto jazyka. Problémem klasifikace stále zůstává fakt, že je velice těžké provést ji v jiných jazycích [19].

5. Prezentace výsledků

Hlavním účelem celé analýzy je překonvertovat nestrukturovaný text do smysluplných informací. Jakmile je analýza dokončena, následuje interpretace dosažených výsledků. Jako nejpřehlednější volba se nabízí interpretace pomocí grafů (obrázek č. 8). Na grafech uvedených níže je k dispozici ukázka způsobu vyjádření dosažených výsledků analýzy. Na

prvním grafu je výsledek znázorněn pomocí procent, naopak u druhého je výsledek uveden jako absolutní hodnota frekvence [19].



Obrázek 8: Ukázka prezentace analýzy sentimentu pomocí grafu

Zdroj:[19]

4.2 Metody analýzy sentimentu

Kritickou částí analýzy sentimentu je výběr vhodné metody. Jako nejčastější využívaná metoda se v literatuře uvádí metoda založená na slovníkové bázi, která využívá „mapovací“ algoritmus, ve kterém program čte text a klasifikuje obsažená slova do různých kategorií, založených na již předdefinovaných kategoriích podle vytvořených slovníků [11]. Dosud nejvhodnějším slovníkem je finanční slovník Loughrana a McDonalda, jelikož je navržen výhradně pro finanční a obchodní sféru [12].

4.3 Slovníky

Běžným slovníkem pro klasifikaci slov je Harvard Psychosociological slovník, neboli Harvard-IV-4 TagNeg (H4N). H4N slovník byl ale kritizován, jelikož je navržen pro oblast sociologie a psychologie. Není možné ho efektivně využít pro oblast financí. Zjistilo se, že téměř ¾ negativních slov podle H4N slovníku nejsou negativní v oblasti financí. Postupem času Loughran a McDonald připravili slovník, který obsahuje 2337 negativních slov speciálně pro finanční sféru. Při svém výzkumu vycházeli z výročních zpráv (formuláře 10-K) z let 1994 – 2008. Jako název pro tento seznam negativních slov použili Fin-Neg [12]. V následující tabulce č. 1 je pro srovnání výčet 10 nejčastěji používaných negativních slov pro celý 10-K dokument.

Tabulka 1: Nejčastější negativní slova z celé 10-K zprávy

CELÝ 10-K DOKUMENT	
10 nejčastěji používaných negativních slov	% všech negativních slov v dokumentu
LOSS	9,73%
LOSSES	5,67%
CLAIMS	3,15%
IMPAIRMENT	3,04%
AGAINST	2,58%
ADVERSE	2,44%
RESTATED	2,09%
ADVERSELY	1,75%
RESTRUCTURING	1,72%
LITIGATION	1,67%

Zdroj: Vlastní zpracování dle [12]

5 SBĚR A PŘÍPRAVA DAT PRO ANALÝZU

Jak již bylo uvedeno výše, tato diplomová práce je zaměřena na hledání sentimentu v mediálních článcích, které čtou investoři. Hlavní otázkou je, jak poté tyto články ovlivňují nebo neovlivňují chování investorů a ceny akcií firem, o kterých je v článcích psáno. Data (články se vztahem k firmám) byla konkrétně sbírána na stránkách Yahoo Finance z mimoburzovního systému Nasdaq a burzy NYSE. Na těchto stránkách jsou volně k dispozici zprávy rozdělené podle firem, kterých se týkají. Klíčovým krokem bylo určit období, za které budou data sbírána. Bylo důležité, aby se jednalo o období, kdy například nebyly velké výkyvy v Dow Jones Economic Sentiment Indicator, zejména kvůli odfiltrování makroekonomického sentimentu. Nakonec bylo vybráno období od 1.2.2016 - 23.2.2016 právě z důvodu stabilního tržního indikátoru sentimentu. Poté co bylo vybráno období, přišel na řadu výběr firem, jejichž články budou sbírány. Z důvodu velkého množství článků za každý sledovaný den se nakonec výběr zúžil jen na 14 největších firem v USA. Pro tyto firmy byl proveden sběr článků za každý den ve sledovaném období. Vzhledem k tomu, že volatilita byla počítána za jednotlivé dny, bylo potřeba veškeré články za každý den pro každou firmu sloučit do jednoho souboru. Tato data byla tedy uložena do textových dokumentů a dále zpracována v programu Statistica (modul Text Miner).

5.1 Vyhodnocení sentimentu

Jak již bylo zmíněno v předešlých kapitolách, nejčastěji se sentiment vyhodnocuje na základě slovníků. I v tomto případě byly proto použity, konkrétně se jedná o slovníky negativních a pozitivních slov od Loughrana a McDonalda [12]. V programu Statistica bylo vyhodnoceno všech 23 dní zvlášť. Pro zajímavost jsou níže uvedeny nejčastěji objevující se slova v článcích za každý den. V tabulce č. 2 jsou nejčastěji se objevující pozitivní slova a v tabulce č. 3 negativní. Víkendy byly pro analýzu sentimentu odebrány, jelikož by výrazně zkreslily výsledek a většina dat byla v tyto dny nedostupná.

Tabulka 2: Vyhodnocení sentimentu - pozitivní slova

Tabulka nejčastějších slov -POZITIVNI					
Datum	Slovo	Četnost	Četnost v člancích	Suma sentimentálních slov	Četnost slova/Suma sent.slov
01.02.2016	strong	36	6	109	33,03%
02.02.2016	strong	21	4	109	19,27%
03.02.2016	good	27	7	123	21,95%
04.02.2016	strong	22	9	123	17,89%
05.02.2016	strong	55	10	110	50,00%
08.02.2016	better	44	10	134	32,84%
09.02.2016	strong	29	8	116	25,00%
10.02.2016	strong	35	8	86	40,70%
11.02.2016	strong	38	10	119	31,93%
12.02.2016	good	47	5	107	43,93%
16.02.2016	strong	30	9	107	28,04%
17.02.2016	great	28	5	108	25,93%
18.02.2016	better	17	6	100	17,00%
19.02.2016	strong	33	8	106	31,13%
22.02.2016	strong	36	12	129	27,91%
23.02.2016	strong	23	9	116	19,83%

Zdroj: Vlastní zpracování

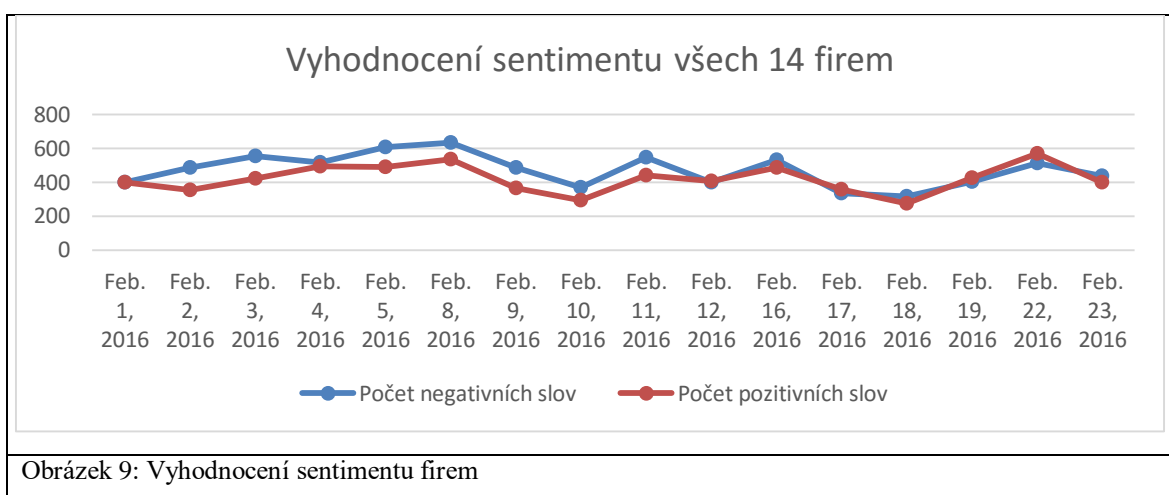
Nejvíce frekventované pozitivní slovo v člancích bylo slovo strong (silný) a poté slovo good (dobrý). Za zmínku stojí také slovo better (lepší), které se v několika člancích objevilo na první pozici.

Tabulka 3: Vyhodnocení sentimentu - negativní slova

Tabulka nejčastějších slov -NEGATIVNI					
Datum	Slovo	Četnost	Četnost v člancích	Suma sentimentálních slov	Četnost slova/Suma sent.slov
01.02.2016	decline	20	8	182	10,99%
02.02.2016	claims	25	4	217	11,52%
03.02.2016	cut	27	4	304	8,88%
04.02.2016	loss	35	7	211	16,59%
05.02.2016	cut	53	7	232	22,84%
08.02.2016	cut	41	9	252	16,27%
09.02.2016	recession	24	7	227	10,57%
10.02.2016	decline	8	6	224	3,57%
11.02.2016	cut	23	7	227	10,13%
12.02.2016	decline	26	7	252	10,32%
16.02.2016	decline	39	9	232	16,81%
17.02.2016	cut	22	5	211	10,43%
18.02.2016	loss	23	5	204	11,27%
19.02.2016	failed	22	2	217	10,14%
22.02.2016	recession	19	4	182	10,44%
23.02.2016	decline	15	4	245	6,12%

Zdroj: Vlastní zpracování

U negativních slov se v článcích objevovalo nejčastěji slovo cut (zkrátit, uštíhnout) a slovo decline (pokles). U negativních slov převažovala celkově větší rozmanitost slov, která se umístila na předních příčkách. Vyhodnocení nebylo tak jednoznačné jako u pozitivních slov. Kromě slov cut a decline lze zmínit i další jako například claim (nárok), loss (zráta) nebo recession (recese). Negativní sentiment se v článcích objevoval téměř dvojnásobně, než tomu bylo u slov pozitivních. V následujícím obrázku č. 9 jsou znázorněny počty pozitivních a negativních slov pro všech 14 firem za každý sledovaný den. V grafu je velmi dobře vidět, zda byl celkový sentiment za určitý den spíše pozitivní nebo negativní. Celkově převažuje negativní sentiment, ale ke konci období se křivky stávají téměř totožnými. Sentiment za jednotlivé firmy je graficky znázorněn v příloze B.



Obrázek 9: Vyhodnocení sentimentu firem

Zdroj: Vlastní zpracování

Poté, co se vyhodnotila pozitivní a negativní slova ze všech článků pro každý sledovaný den, bylo ještě potřeba upravit počet pozitivních slov. Program Statistica, který pozitivní slova vyhledával, mohl za pozitivní slovo považovat i takové, jehož význam je negativní. Tato chyba může nastat v případě, že pozitivní slovo má předložku not. Statistica takto napsané slovo vyhodnotí pozitivně, i když je očividné, že slovo je negativního rázu. Aby se této chybě zamezilo, bylo ještě potřeba upravit slovník pozitivních slov tak, že se před všechna slova přidala předložka not a takto upravený slovník byl použit pro odhalení těchto slov. Vyhledaná slova se musela odečíst od pozitivních a přičíst k negativním. Těchto slov bylo v článcích pouze zanedbatelné množství. Po úpravě všech negativních a pozitivních slov byl vypočítán tón zpráv jako podíl negativních slov odečtených od pozitivních a jejich součtu. Příprava dat pro následnou regresi je vždy nejsložitější a časově nejnáročnější část v celém procesu analýzy sentimentu.

5.2 Ukazatele pro predikci volatility

Před samotným začátkem regresní analýzy byla získána data pro všechny nezávislé proměnné ze stránek Yahoo Finance. Vybrány byly ukazatele, které se v minulosti ukázaly jako významné determinanty budoucí volatility [18]. Ty pak byly využity pro predikci závislé proměnné budoucí 3-denní volatility (vypočtená za dny t , $t+1$ a $t+2$). V publikaci od Loughrana a McDonalda [12] bylo pro výpočet volatility zvoleno období za tři dny, proto bylo toto období vybráno i zde. Pro delší období, za které by byla volatilita měřena, by bylo také potřeba zvolit delší časové období.

Jako nezávislé proměnné byly vybrány:

1. Tržní kapitalizace za den $t-1$
2. P/E za den $t-1$
3. Volume (objem obchodů) za den $t-1$
4. Tón zpráv za den $t-1$
5. Délka článků za den $t-1$
6. Negativní slova/délka článků za den $t-1$
7. Pozitivní slova/délka článků za den $t-1$
8. Typ burzy (Nasdaq/NYSE) viz tabulka č. 4
9. Historická 3-denní volatilita (ve dnech $t-3$, $t-2$ a $t-1$)

Tabulka 4: Základní statistiky proměnné Typ burzy

Typ burzy	Označení burzy	Suma výskytu	Počet firem
Nasdaq	1	64	4
NYSE	0	160	10

Zdroj: Vlastní zpracování

5.2.1 Tržní kapitalizace

Tržní kapitalizace vyjadřuje hodnotu podniku na základě tržního ohodnocení cen všech jeho akcií. Hodnota podniku je získána součinem všech vydaných akcií a jejich aktuální tržní hodnoty. S rostoucí tržní kapitalizací je často spojen vyšší obrat akcií. To znamená, že se akcie více prodávají a kupují. Z toho důvodu používají investoři často tento ukazatel jako měřítko, aby vyhodnotili objem obchodu a tržní likviditu podniku [6]:

$$\text{Tržní kapitalizace} = \text{součet vydaných akcií} * \text{aktuální tržní hodnota akcií.} \quad (7)$$

5.2.2 P/E

P/E je ukazatel, který označuje, kolik jsou ochotni akcionáři zaplatit za 1 Kč zisku na akcii. Případně také odhaduje počet let potřebných ke splacení ceny akcie jejím výnosem (při předpokladu konstantní výše zisku a ceny). Relativně vysoká hodnota v rámci odvětví znamená, že investoři očekávají v budoucnu vysoký růst dividend nebo že akcie obsahuje malé riziko, díky kterému se investoři spokojí s menším výnosem [21]. Naopak nízká hodnota může znamenat nízký růstový potenciál firmy či vyšší rizikovost firmy. V praxi ukazatel odráží hodnocení budoucího vývoje firmy investory.

$$P/E = \text{tržní cena akcie} / \text{čistý zisk na akcii}. \quad (8)$$

Výslednou hodnotu P/E mohou zkreslit některé faktory [20]:

- použité účetní metody – velikost zisku firem závisí i na použitých účetních metodách, které se mohou lišit v jednotlivých zemích,
- jednorázové obchodní a finanční operace – především prodej části podniku, mimořádné odpisy apod.,
- aktuálnost dosazovaných údajů – u tržní ceny akcie se mohou vyskytnout pochybnosti pouze u veřejně neobchodovatelných akcií nebo u akcií s velmi nízkou likviditou, větší problém je s aktuálností čistého zisku na akcii - největší riziko nastává na začátku účetního období (před účetní závěrkou), kdy podniky využívají zastaralé údaje z předminulého roku.

5.2.3 Volume (objem obchodů)

Volume (často uváděn pod zkratkou VOL), nebo také česky objem obchodů, je indikátor vyjadřující celkové množství kontraktů zobchodovaných v rámci specifického časového úseku (např. hodina, den, týden, měsíc) [20]. Jinými slovy, hodnota volume říká, jak hodně se v daný okamžik obchodovalo - vysoká hodnota VOL znamená, že v daný časový úsek proběhlo mnoho obchodů a naopak. Volume jako nezávislá proměnná byla vybrána, jelikož úzce souvisí s volatilitou. Pokud je objem obchodů vysoký, dá se očekávat i vyšší volatilita, jelikož je trh více v pohybu.

6 PREDIKCE VOLATILITY POMOCÍ REGRESNÍHO MODELU

V regresní analýze se obecně analyzuje vztah mezi jednou proměnnou zvanou cílová nebo závislá proměnná, a několika dalšími, které se nazývají nezávislé nebo ovlivňující proměnné. Vztah se reprezentuje matematickým modelem, což je rovnice, jež svazuje nezávislé a závislé proměnné a pravděpodobnostní předpoklady, které by měl vztah splňovat. Závisle proměnná je spojena s nezávisle proměnnými funkcí nazývanou regresní funkce, jež obsahuje několik neznámých parametrů. Jestliže tato funkce je lineární v těchto parametrech, mluví se o lineárním regresním modelu. Statistické problémy, které jsou důležité v regresní analýze, jsou [3]:

- získání statistických odhadů neznámých parametrů regresní funkce;
- testování hypotéz o těchto parametrech;
- ověřování předpokladů regresního modelu.

Jestliže graf ukáže lineární vztah mezi proměnnými, usiluje se o zachycení vztahu tím, že se body proloží přímkou. Stane se to hlavně tehdy, jestliže je korelace mezi proměnnými menší. Hledá se tedy přímka, jež je experimentálními bodům co možná nejbliže. Z tohoto požadavku plyne, že přímka by měla být nejbliže k bodům ve vertikálním směru, protože chyby predikce se týkají závislých hodnot.

Rozdílu mezi naměřenou a predikovanou hodnotou říkáme reziduální hodnota predikce nebo chyba predikce a značíme ji symbolem e . Dobře proložená přímka má tvar

$$y = a + b \tag{9}$$

a minimalizuje velikosti reziduálních hodnot po hodnoty (x_i, y_i) , kterými se přímka prokládá. Je mnoho způsobů, jak to provést. Nejčastěji se používá metoda nejmenších čtverců.

Hodnoty parametrů a, b přímky $y = a + bx$ lze získat metodou nejmenších čtverců tak, aby byl minimální součet druhých mocnin reziduálních hodnot

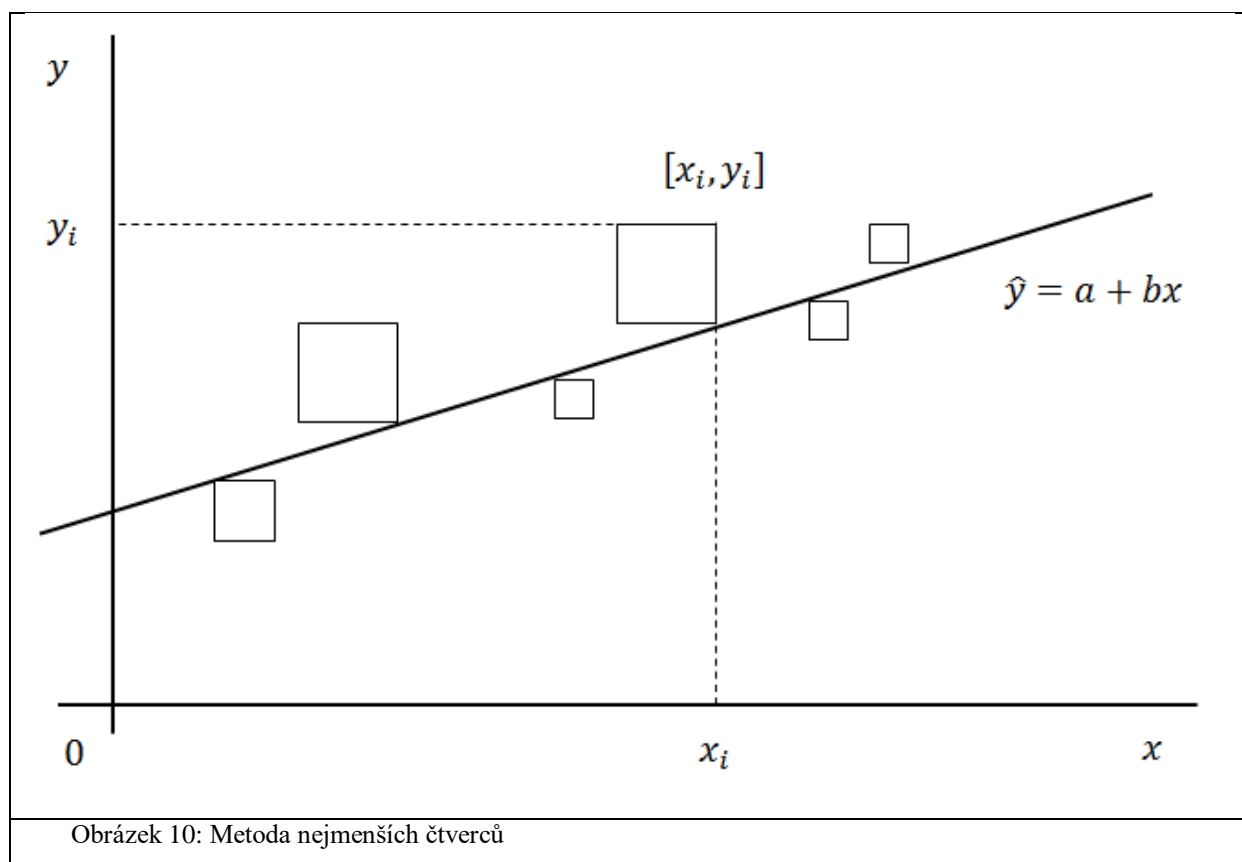
$$s_r^2 = \sum e_i^2 = \sum (y_i - a - bx_i)^2 \tag{10}$$

vzhledem k parametrům a, b . Grafickou interpretaci odchylek od regresní přímky pro metodu nejmenších čtverců ukazuje obrázek č. 10. Minimalizují se sečtené čtverce úseček, které

vyznačují vzdálenost bodu od proložené přímky ve směru osy y . Výpočet tohoto minima vede k optimálním hodnotám

$$b = r \frac{s_y}{s_x}, \quad a = \bar{y} - b\bar{x}, \quad (11)$$

kde r je korelace obou proměnných a s_x, s_y jsou směrodatné odchylky naměřených hodnot x a y . Nalezená přímka se nazývá regresní přímka. Hodnota \hat{y}_i je odhad cílové proměnné pomocí regresního vztahu $\hat{y}_i = a + bx_i$ [3].



Obrázek 10: Metoda nejmenších čtverců

Zdroj:[3]

Existují různé typy regresních modelů podle tvaru regresní funkce. Lze rozlišit základní typy:

- a) modely lineární vzhledem k parametrům (např. regresní přímka, regresní parabola, regresní hyperbola apod.),
- b) modely nelineární vzhledem k parametrům, které je však možné transformovat na lineární (např. regresní mocninná funkce, regresní exponenciální funkce),
- c) modely nelineární, které se nedají jednoduše transformovat na lineární tvar.

U prvních dvou typů se odhady parametrů provádějí nejčastěji metodou nejmenších čtverců. U posledního typu tato metoda není vhodná a odhady se provádějí jinými metodami (např. metodou částečných součtů, metodou dílčích průměrů atd.).

Jednoduchý model lineární regrese se hodí např. pro vztah výšky a hmotnosti člověka – tedy pro jednu nezávislou veličinu X . V této práci je však zkoumán vliv sentimentu a ostatních faktorů na změnu volatility, proto je třeba využít vícerozměrný model lineární regrese [3].

6.1 Vícerozměrná lineární regrese

V analýze založené na vícerozměrné regresi se hledají hodnoty závisle proměnné z lineární kombinace hodnot několika nezávisle proměnných. Vícerozměrný model lineární regrese je následující [3]:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + E_i, \quad (12)$$

kde $i = 1, 2, \dots, n$, β_i jsou neznámé parametry a E_i jsou nezávislé náhodné veličiny, mající stejné rozdělení pravděpodobnosti se střední hodnotou 0 a kovarianční maticí $\sigma^2 \mathbf{I}$, přičemž \mathbf{I} je jednotková matice. Parametry $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ se odhadují metodou nejmenších čtverců. Funkce $\hat{Y} = B_0 + B_1 x_1 + B_2 x_2 + \dots + B_k x_k$ se nazývá výběrová regresní funkce.

Poté, co byly vybrány všechny nezávislé proměnné X a závislá proměnná Y a byl vybrán vhodný model mnohonásobné regrese, byla provedena v programu Statistica mnohonásobná regrese. Než byla regrese provedena, bylo zapotřebí zjistit, zda se určité nezávislé proměnné nekorelují. V tabulce č. 5 je znázorněna korelační matice všech proměnných. Z výsledku je vidět, že silně korelují proměnné tón zpráv a pozitivní slova/délka zpráv a negativní slova/délka zpráv. Pokud by tyto proměnné byly ponechány, byly by výsledky regresního modelu zkreslené. Pro další analýzu byla proto proměnná tón zpráv vyřazena.

Poslední úprava dat byla provedena na proměnných tržní kapitalizace a volume, které zlogaritmovány. Důvodem bylo jejich odchýlení od normálního rozdělení. Po logaritmizaci se rozdělení těchto proměnných přiblížilo k normálnímu.

Tabulka 5: Korelace nezávislých proměnných

Variable	Correlations (Tabulka1 - pro regresi-1)								
	log market cap	log volume	tón zpráv	délka zpráv	positive/délka zpráv	negative/délka zpráv	typ burzy	vol. minulá	vol. budoucí
log market cap	1,000000	0,097478	-0,047187	0,261582	-0,079505	-0,040797	-0,154639	-0,343361	-0,417250
log volume	0,097478	1,000000	-0,181906	0,298221	-0,249568	0,036478	0,009311	0,350977	0,260474
tón zpráv	-0,047187	-0,181906	1,000000	0,068123	0,589839	-0,730932	0,031734	-0,111138	-0,222374
délka zpráv (slov)	0,261582	0,298221	0,068123	1,000000	-0,167125	-0,289002	0,170592	-0,064729	-0,039188
positive/délka zpráv	-0,079505	-0,249568	0,589839	-0,167125	1,000000	-0,015530	-0,022516	-0,072366	-0,157282
negative/délka zpráv	-0,040797	0,036478	-0,730932	-0,289002	-0,015530	1,000000	-0,111499	0,121375	0,187150
typ burzy	-0,154639	0,009311	0,031734	0,170592	-0,022516	-0,111499	1,000000	-0,015723	-0,039602
vol. minulá	-0,343361	0,350977	-0,111138	-0,064729	-0,072366	0,121375	-0,015723	1,000000	0,365213
volatilita budoucí	-0,417250	0,260474	-0,222374	-0,039188	-0,157282	0,187150	-0,039602	0,365213	1,000000

Zdroj: Vlastní zpracování v programu Statistica

6.2 Výsledky vícerozměrné regrese

V první regresi byla jako závislá proměnná použita 3-denní budoucí volatilita⁷, která následovala bezprostředně další dny po volatilitě historické, která byla použita jako nezávislá proměnná. Jak je vidět z výsledné tabulky č. 6, na hladině významnosti $\alpha=0,05$, byla jako statisticky významná proměnná vyhodnocena tržní kapitalizace, která se znaménkem mínus snižuje velikost volatility. Další významné proměnné jsou typ burzy a volume. Typ burzy Nasdaq snižuje volatilitu a volume zvyšuje. Sentiment v tomto případě hraje nevýznamnou roli. Z výsledku je alespoň vidět, že negativní slova působí na zvýšení volatility a naopak slova pozitivní volatilitu snižují. Další hodnota, která je na modelu důležitá je hodnota R^2 , neboli koeficient determinace. V tomto případě hodnota okolo 0,36 signalizuje, že model vysvětlil asi 36 % rozptylu v datech. To naznačuje, že na predikci volatility mají silný vliv i jiné faktory, které v této práci nebyly uvažovány. Přesto výsledky mnohonásobné regrese nejsou zanedbatelné a regresní model poskytuje statisticky významnou predikci.

Jelikož z předchozího modelu nevyplýval fakt, že by sentiment významně ovlivnil budoucí volatilitu, pro další regresní model (regrese 2) byla jako budoucí 3-denní volatilita zvolena data, která nenásledovala bezprostředně další dny, tj. za dny $t+3$, $t+4$ a $t+5$. Jedná se tedy o predikci volatility s větším časovým rozmezím, jelikož je možné, že investoři mohou reagovat na sentiment v člancích s delším časovým odstupem.

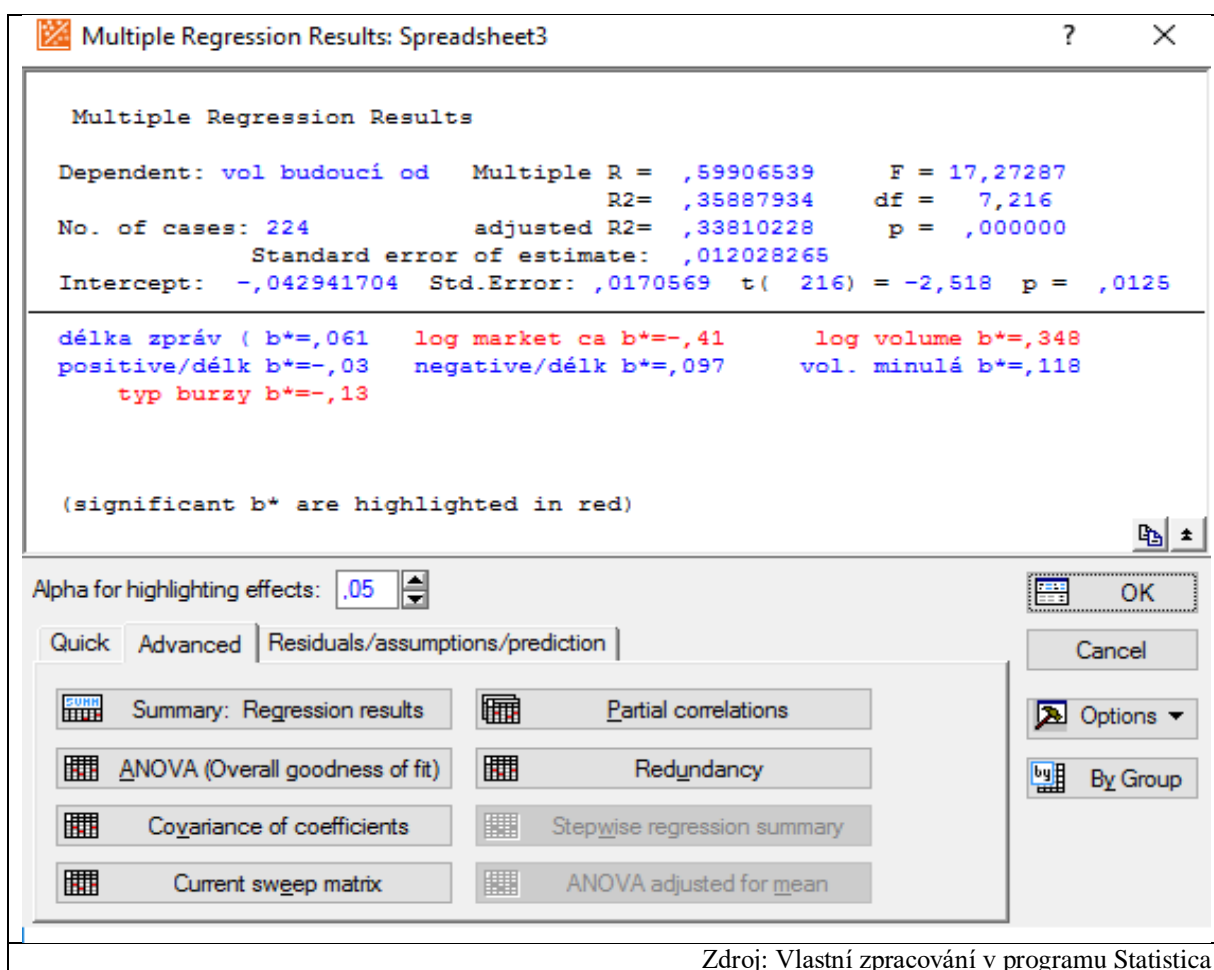
⁷ Výpočet podle vzorce:
$$\sqrt{\frac{\sum((r_i)_{N-1} - (\text{průměr}(r_i)))^2}{n-1}}$$

kde r_i je výnos za den,

n je počet dní, za které se volatilita sleduje,

N je den začátku měření volatility.

Tabulka 6: Výsledky vícerozměrné regrese 1



Po provedené regresi z tabulky č. 7 je opravdu vidět, že nyní se sentiment stává významným determinanem v predikci volatility. Opět zůstal zachován předpoklad, že pozitivní slova by měla volatilitu snižovat a naopak slova negativní zvyšovat. Spolehlivost modelu se od předchozího snížila, což se dalo vzhledem k delšímu horizontu predikce očekávat.

V další tabulce č. 8 jsou zobrazeny základní statistiky všech vybraných proměnných.

Tabulka 7: Výsledky vícerozměrné regrese 2

Multiple Regression Results: Tabulka1 - pro regresi-1
? ×

Multiple Regression Results

Dependent: volatilita bud	Multiple R = ,57176000	F = 14,98683
	R2= ,32690949	df = 7,216
No. of cases: 224	adjusted R2= ,30509637	p = ,000000
Standard error of estimate: ,011031381		
Intercept: -,001312608	Std.Error: ,0156433	t(216) = -,0839 p = ,9332

log market ca b*=-,43	log volume b*=?,208	délka zpráv (b*=?,061
positive/délk b*=-,12	negative/délk b*=?,152	vol. minulé b*=?,119
typ burzy b*=-,10		

(significant b* are highlighted in red)

Alpha for highlighting effects: .05	OK
Quick Advanced Residuals/assumptions/prediction	Cancel
Summary: Regression results	Options ▼
	By Group

Zdroj: Vlastní zpracování v programu Statistica

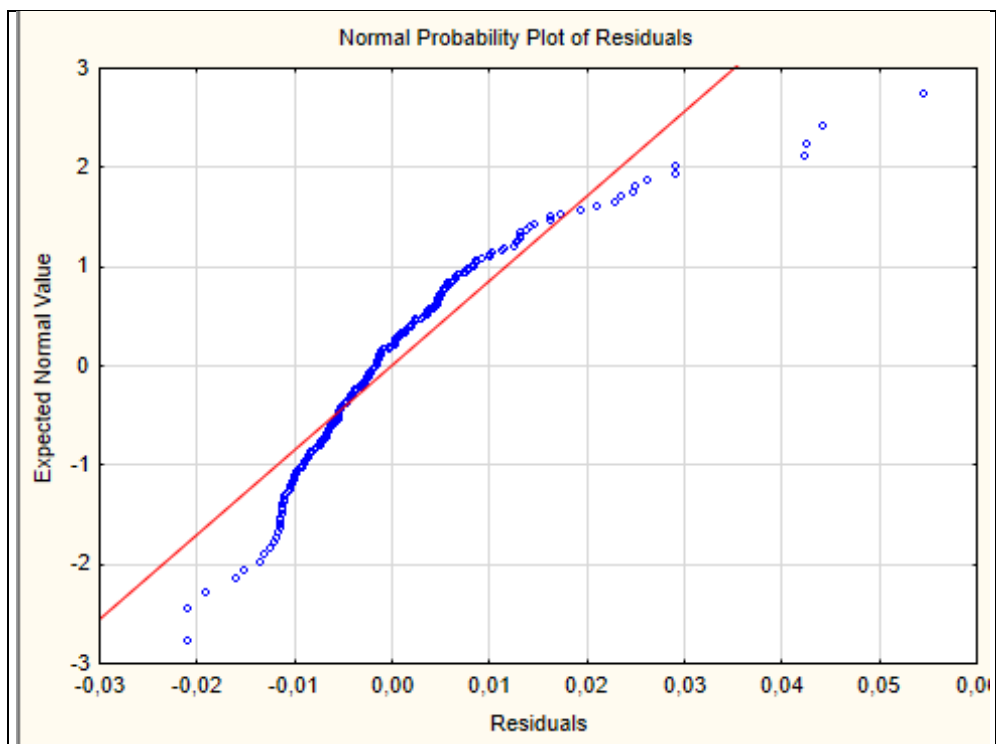
Tabulka 8: Základní statistiky vybraných proměnných

Variable	Descriptive Statistics (Tabulka1 - pro regresí-1)				
	Valid N	Mean	Minimum	Maximum	Std.Dev.
log market cap	224	2,11	1,4140	2,7	0,39
log volume	224	7,17	6,2673	8,0	0,39
tón zpráv	224	-0,06	-0,8947	0,8	0,33
volatilita budoucí	224	0,02	0,0017	0,1	0,01
délka zpráv (slov)	224	17018,37	612,0000	107895,0	13696,67
positive/délka zpráv	224	0,00	0,0002	0,0	0,00
negative/délka zpráv	224	0,00	0,0002	0,0	0,00
typ burzy	224	0,29	0,0000	1,0	0,45
vol. minulá	224	0,02	0,0017	0,1	0,02

Zdroj: Vlastní zpracování v programu Statistica

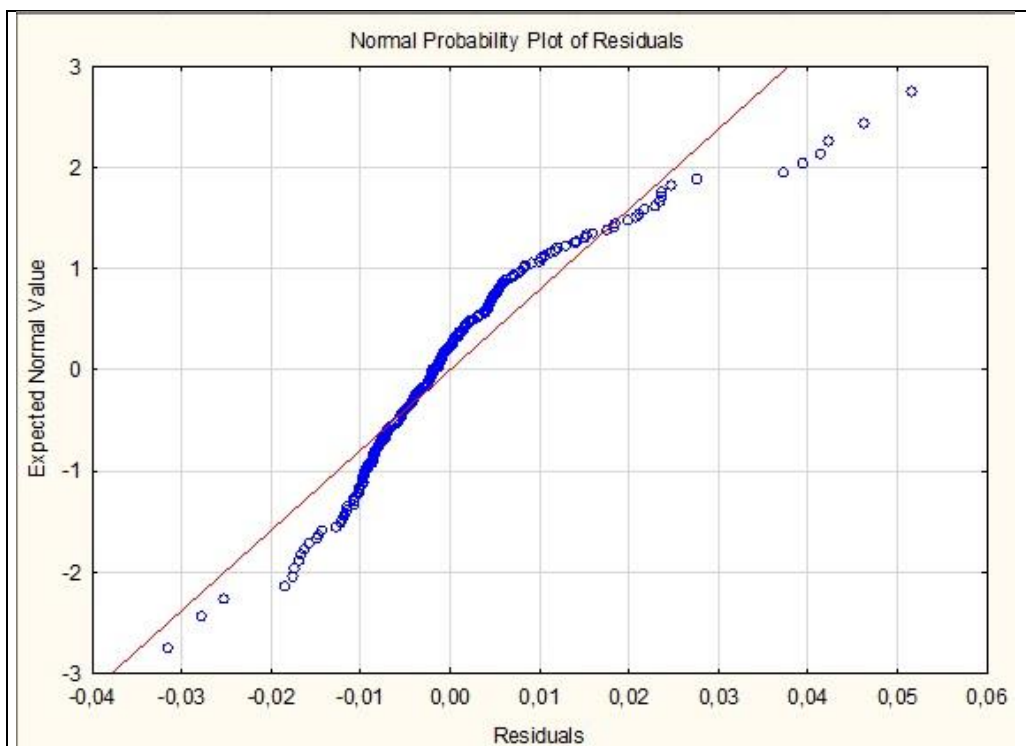
6.3 Ověření předpokladů modelu

Posledním bodem při tvorbě regresního modelu je ověření všech předpokladů, o kterých byla řeč v předchozích kapitolách. Jako první byla ověřena normalita reziduí. Normalita může být ověřena testováním normality odchylek modelu od skutečnosti (testováním normality reziduí). K ověření normality na základě reziduí slouží reziduální analýza. Zde byla vybrána jako možnost zobrazení normality pomocí normálního p -grafu reziduí. Aby byl předpoklad normality splněn, všechny body by se měly pohybovat na přímce nebo velmi blízko přímky. Jak ukazuje obrázek č. 11 a 12, předpoklad normality není zcela splněn, jelikož se body nechovají úplně jako přímka. Výsledek je však blízko normálnímu rozdělení.



Obrázek 11: Normální p-graf reziduí 1. regrese

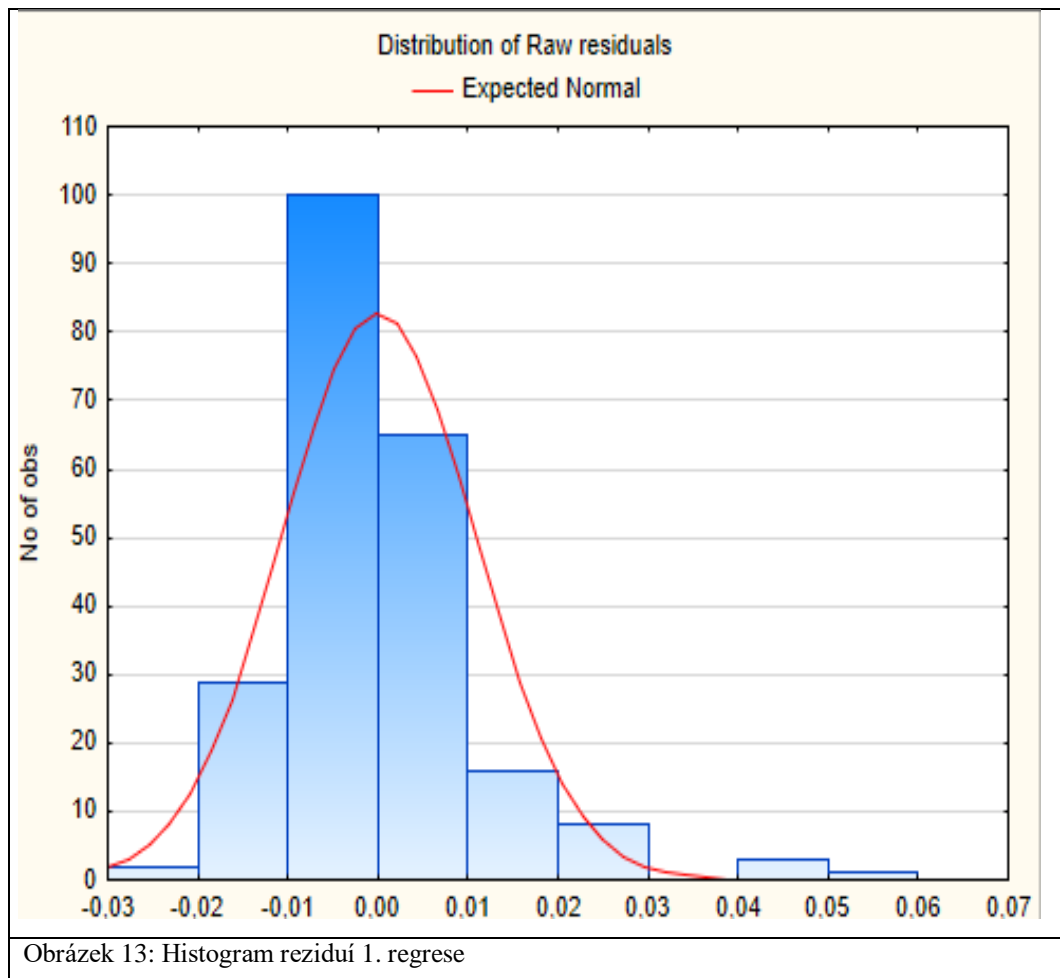
Zdroj: Vlastní zpracování v programu Statistica



Obrázek 12: Normální p-graf reziduí 2. regrese

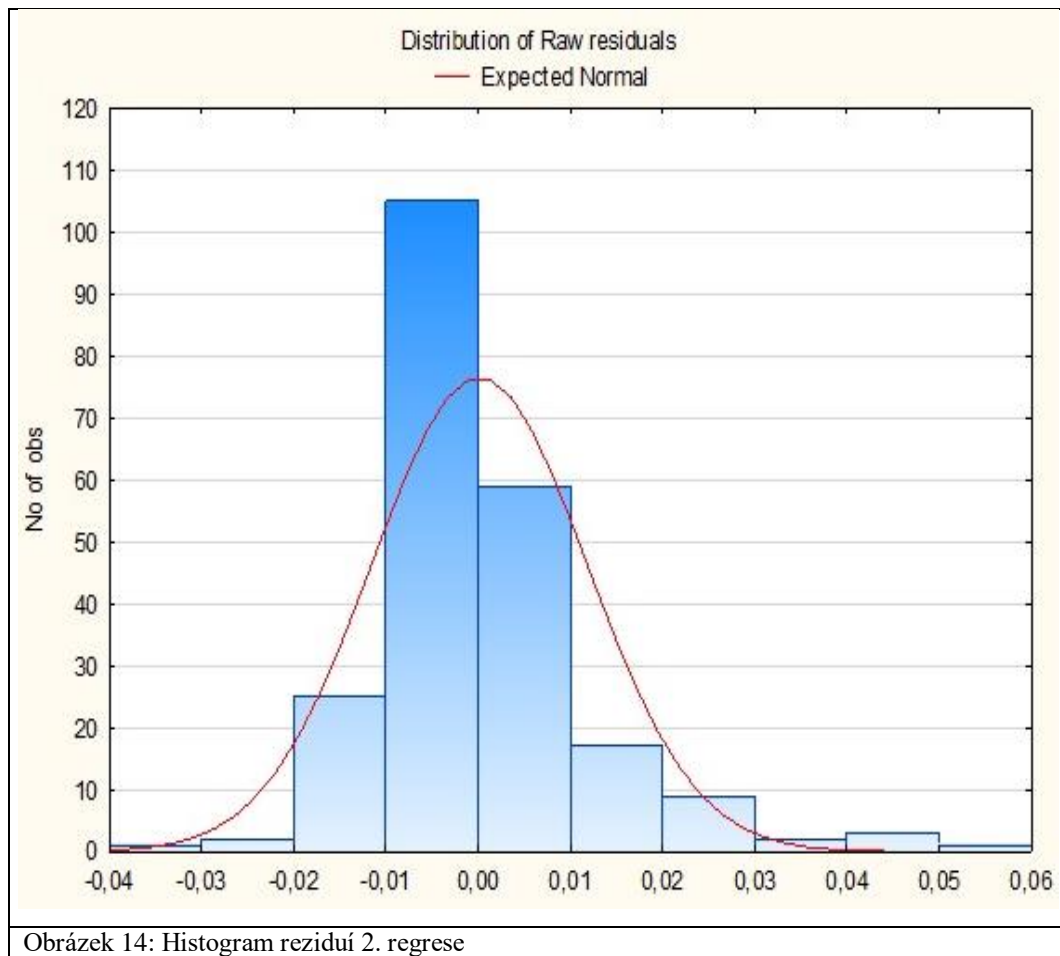
Zdroj: Vlastní zpracování v programu Statistica

Rezidua mohou být znázorněna také histogramem reziduí, kde je z obrázků č. 13 a 14 opět vidět, že data nejsou zcela z normálního rozdělení. Můžeme vidět mírné zešikmení zprava. Po odebrání odlehlých hodnot by se mohlo zešikmení zmírnit. Opět je ale zřejmé, že rozdělení reziduí je blízko normálnímu rozdělení.



Obrázek 13: Histogram reziduí 1. regrese

Zdroj: Vlastní zpracování v programu Statistica

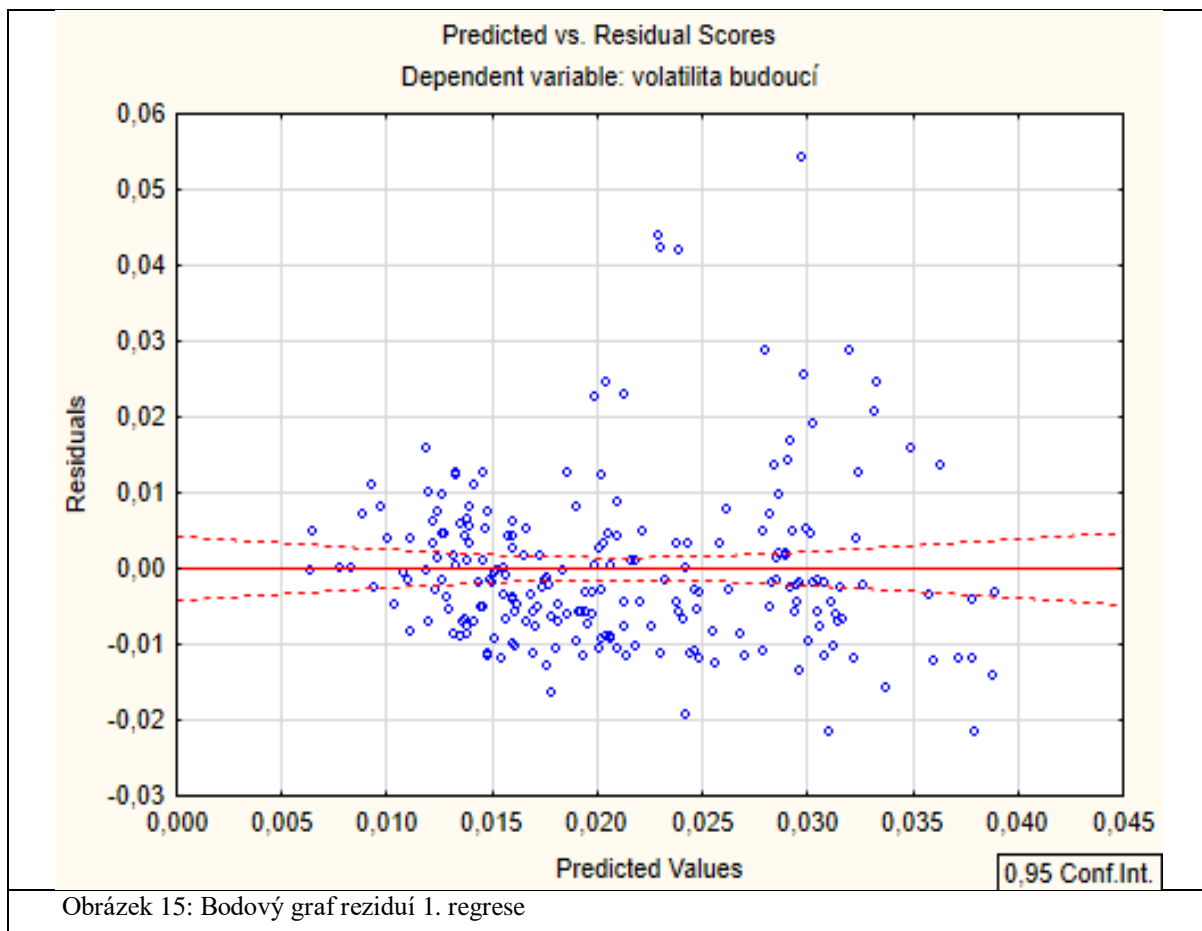


Obrázek 14: Histogram reziduí 2. regrese

Zdroj: Vlastní zpracování v programu Statistica

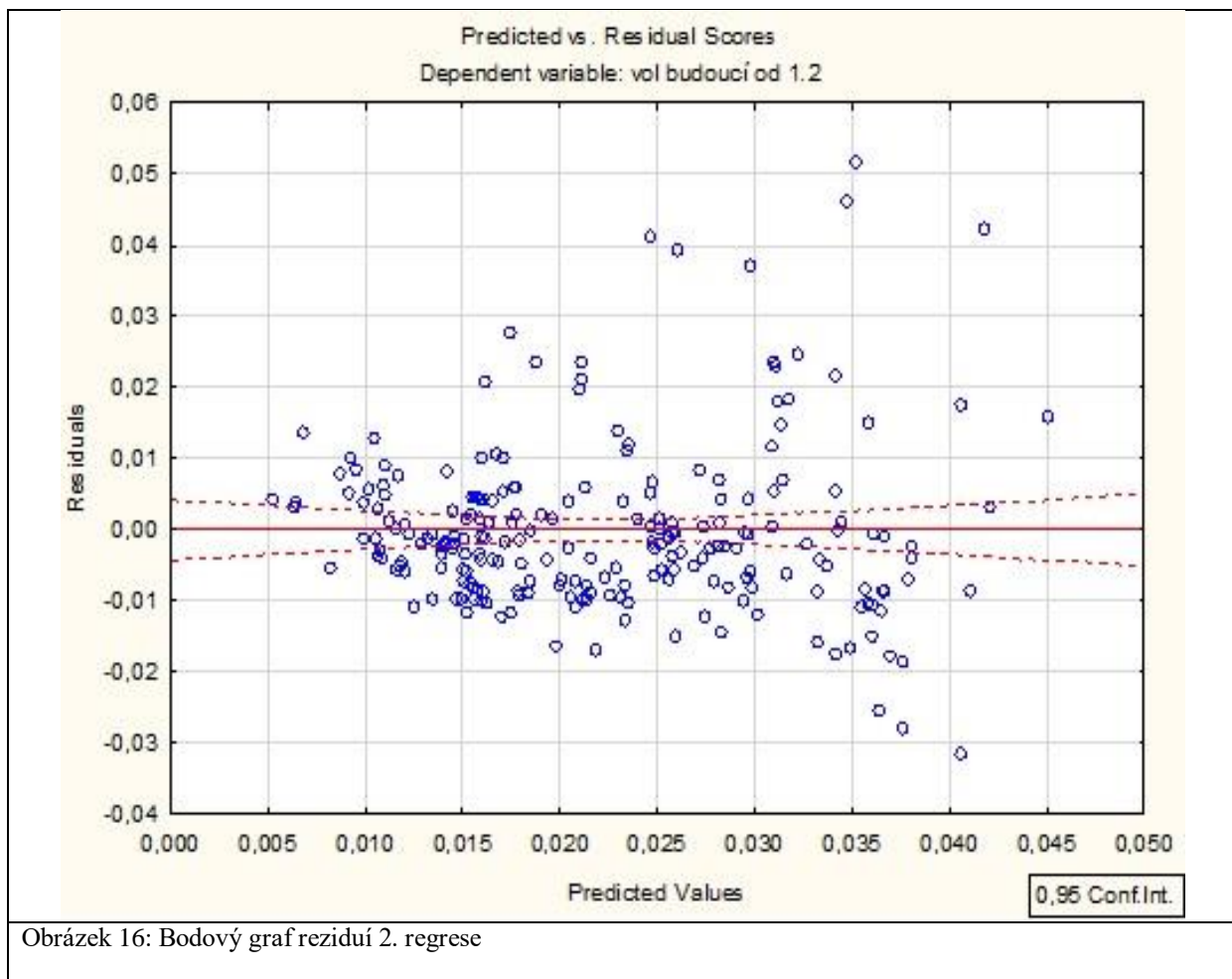
Dalším předpokladem regresního modelu je, že střední hodnota chybové složky je rovna 0 a chybová složka má konstantní rozptyl (homoskedasticita). Předpoklad homoskedasticity lze ověřit pomocí Scatter plotu (bodový graf). Bodový graf zobrazuje řadu jako sadu bodů. Hodnoty jsou vyjádřeny polohou bodů v grafu. Kategorie jsou vyjádřeny různými značkami v grafu. Bodové grafy jsou obvykle používány ke srovnání agregovaných dat různých kategorií.

Z výsledného obrázku č. 15 a 16 lze soudit, že rezidua mají konstantní rozptyl, jelikož se pohybují až na pár odlehlých bodů ve stejných mezích, které představuje pomyslná přímka. Rezidua modelu jsou zhruba stejně rozptýlena kolem nulové střední hodnoty. Správnost modelu se dá vysvětlit i tak, že po celé délce x-ové osy, jsou rezidua rovnoměrně zastoupena jak nad nulou, tak i pod ní.



Zdroj: Vlastní zpracování v programu Statistica

Jako další předpoklad, který byl splněn již před začátkem regrese, byla nekorelovanost všech nezávislých proměnných.



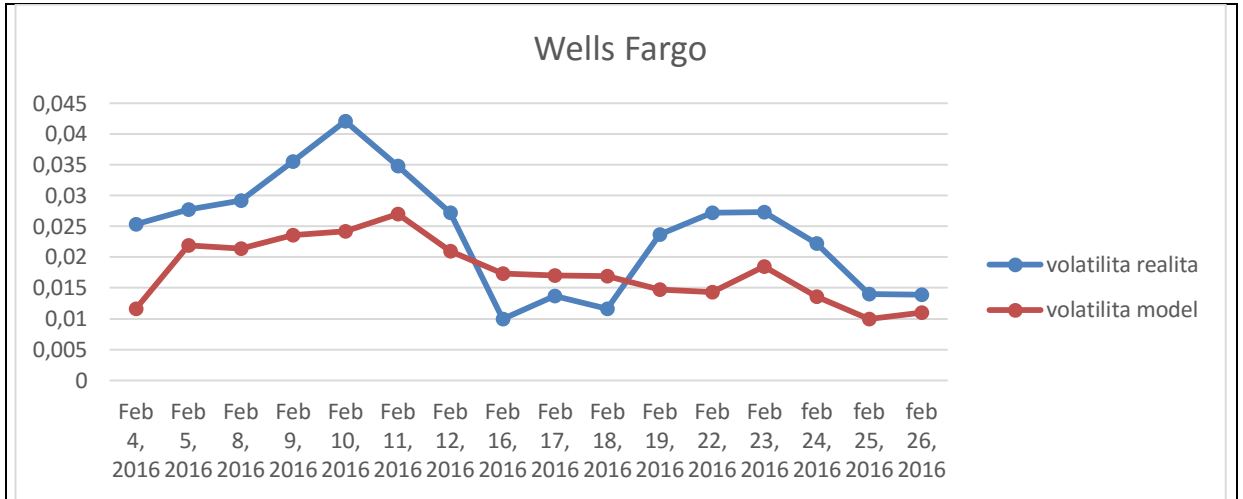
Obrázek 16: Bodový graf reziduí 2. regrese

Zdroj: Vlastní zpracování v programu Statistica

6.4 Vizualizace predikované volatility

V poslední části analýzy bude pozornost zaměřena na vizualizaci předpovězené volatility pomocí grafu. Po dosažení regresních koeficientů do vzorce, který byl zmíněn již na začátku této kapitoly, lze obdržet predikované hodnoty \hat{y}_i . Rozdíly mezi závislou proměnnou, neboli 3-denní volatilitou, která byla do regresního modelu vložena (očekávaná hodnota) a predikovanou volatilitou, jsou právě rezidua modelu od skutečnosti. Jak již bylo zmíněno, hodnota koeficientu determinace R^2 se u modelů pohybuje mezi 0,32 a 0,36 a je tedy jasné, že na budoucí volatilitu působí mnohem více faktorů, které ji ovlivňují, a které nebyly v této analýze brány v úvahu. Následující grafy jsou tedy vizuální ukázkou dosažených predikcí. Samozřejmě se vzhledem k uvedeným hodnotám R^2 nedá očekávat, že křivky reálné volatility a volatility predikované spolu budou zcela korelovat.

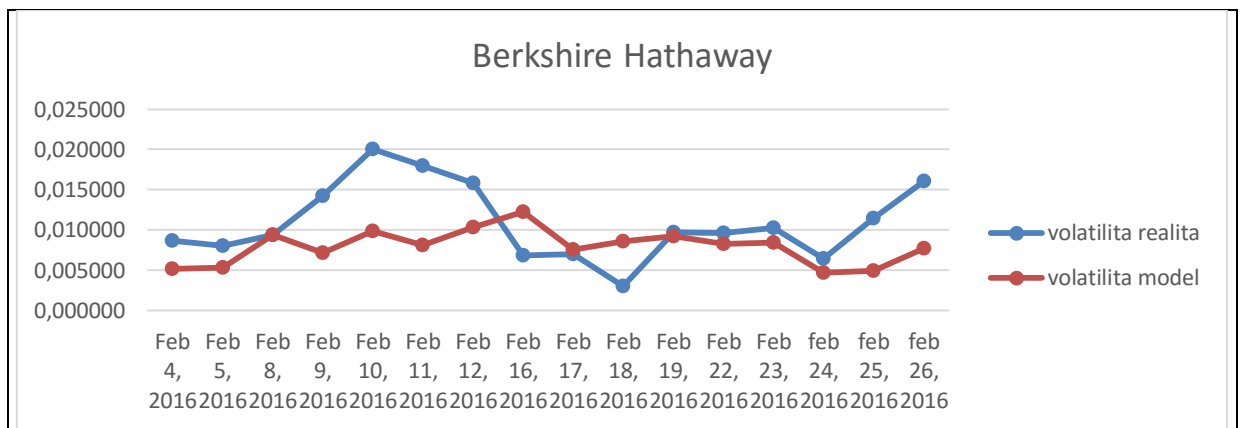
Jako první byla vybrána firma Wells Fargo, kde je zřejmý společný trend křivek (obrázek č. 17). Reálná volatilita se vychyluje více než predikovaná, ale dalo by se konstatovat, že pokud reální volatilita klesala, model vícerozměrné regrese předpověděl také klesající trend volatility a naopak.



Obrázek 17: Porovnání volatilit model vs. realita Wells Fargo

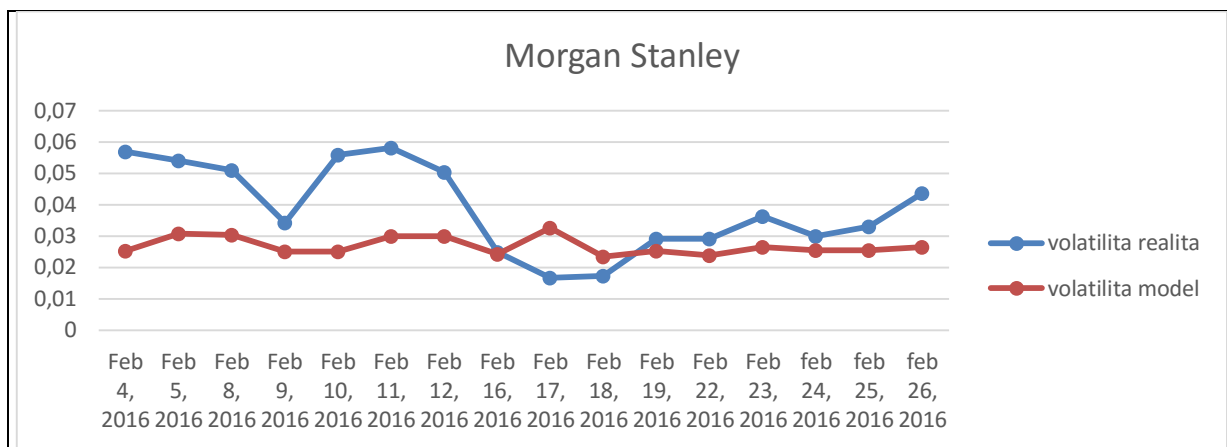
Zdroj: Vlastní zpracování

Následující dva obrázky č. 18 a 19 firem Berkshire Hathaway a Morgan Stanley zachycují podobné výsledky. Lze vidět náznaky společného trendu, ale skutečné vychýlení volatilit bylo větší. Je tedy jasné, že vychýlení způsobily i jiné faktory.



Obrázek 18: Porovnání volatilit model vs. realita Berkshire Hathaway

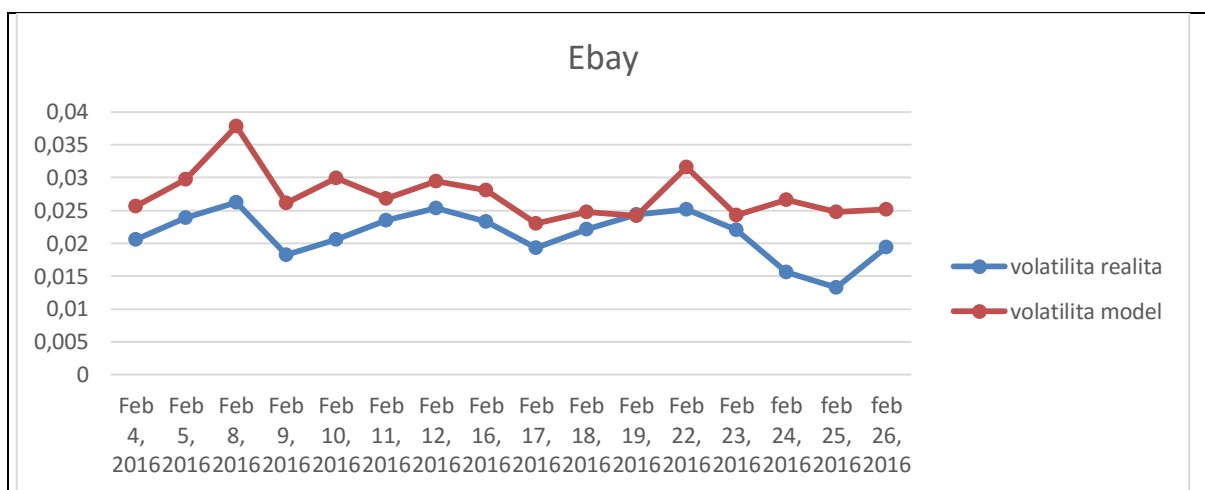
Zdroj: Vlastní zpracování



Obrázek 19: Porovnání volatilit model vs. realita Morgan Stanley

Zdroj: Vlastní zpracování

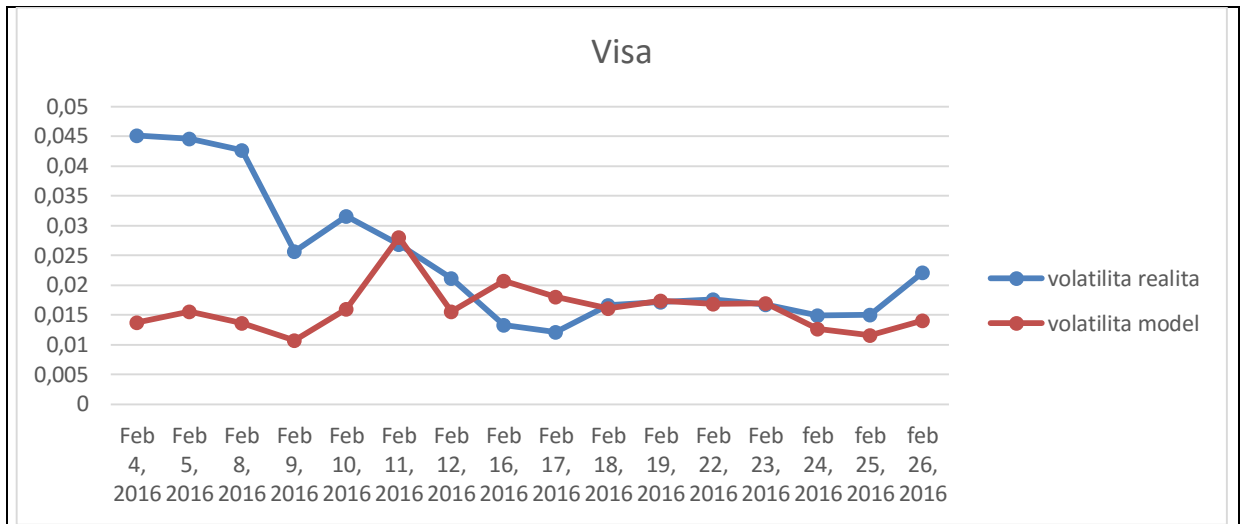
Další obrázek č. 20 firmy Ebay mírně kopíruje trend reálné volatilit, model ale předpověděl za celé sledované období vyšší volatilitu, než ve skutečnosti byla.



Obrázek 20: Porovnání volatilit model vs. realita Ebay

Zdroj: Vlastní zpracování

Poslední uvedený obrázek č. 21 firmy Visa ukazuje shodu modelu s realitou v 5 bodech. V dalších bodech se model neshoduje, ale opět se dá říci, že mírně kopíruje trend reálné volatilit. Zbytek firem viz. příloha C.



Obrázek 21: Porovnání volatilit model vs. realita Visa

Zdroj: Vlastní zpracování

ZÁVĚR

Cílem této práce bylo predikovat volatilitu vybraných kapitálových firem pomocí analýzy sentimentu. Pro splnění cíle bylo vybráno 14 největších amerických firem obchodujících na burze NYSE a na mimoburzovním trhu Nasdaq. Dalším krokem bylo posbírat všechny články, které byly o těchto firmách vydány různými americkými médii za vybrané období 1.2. – 23.2.2016. Všechny články za každý den jsou dohledatelné na Finance Yahoo. Články byly dále ukládány do textového dokumentu. Sentiment v textech se objevuje v různých odvětvích a v každém mohou mít různá slova odlišný význam. Pro správné odhalení sentimentu v textech bylo nutné pracovat se správným slovníkem, který je zaměřen na finanční sféru. Pro vyhodnocení sentimentu byl vybrán slovník Loughrama a McDonalda pozitivních a negativních slov. Všechny články byly následně zpracovány v programu Statistica pomocí modulu Text mining a vybraných slovníků. Výsledná data byla zpracována, aby se dala vhodně použít jako nezávislá proměnná pro predikci volatility. Nejdříve byla vyhodnocena četnost pozitivních a negativních slov v článcích a poté byly četnosti převedeny do poměru k délce článků. Jako další ukazatel byl vypočítán také tón zpráv.

K ukazatelům, které byly vypočítány z analýzy sentimentu, byly přidány i další – tržní kapitalizace, P/E, délka článků, historická volatilita, volume (objem obchodů) a typ burzy. K odhalení závislosti volatility na sentiment v článcích byla použita regresní analýza, konkrétně vícerozměrná regresní analýza. Jako nezávislé proměnné byly zadány všechny hodnoty sentimentu a ostatní ukazatelé. Ještě před začátkem vícerozměrné regrese byla sestavena korelační matice všech nezávislých proměnných. Z matice bylo zjištěno, že se významně korelují tón zpráv s poměrem pozitivních slov a délky články a také s poměrem negativních slov a délky článků. Pro další použití byla proměnná tón zpráv odebrána. Následovala vícerozměrná regrese. Jako závislá proměnná byla vybrána 3-denní volatilita od 1.2.2016. Za historickou volatilitu bylo vybráno období o tři dny dříve. Z výsledné tabulky bylo zjištěno, že sentiment na hladině významnosti $\alpha=0,05$ významně neovlivnil predikovanou volatilitu a jako významné proměnné byly identifikovány typ burzy, volume a tržní kapitalizace. Volume dle modelu vícerozměrné regrese ovlivňuje budoucí volatilitu kladně a ostatní významné proměnné volatilitu ovlivňují záporně.

Jelikož nebyl z této regrese zjištěn významný vliv sentimentu, regrese byla provedena znovu. Nyní byla pozměněna budoucí 3-denní volatilita o další tři dny a znovu se zkoumal vliv sentimentu a dalších ukazatelů na výsledek modelu. Ve druhém regresním modelu byly výsledky z pohledu analýzy sentimentu zajímavější. Z výsledných hodnot vícerozměrné

regrese se sentiment stává významným v predikci volatility. Výsledné hodnoty korelují s předpokladem, že pozitivní sentiment v člancích snižuje výši volatility. Naopak negativní sentiment volatilitu zvyšuje. Koeficient determinace R^2 byl ale pouze 0,32. Z tohoto výsledku se dá tvrdit, že by bylo vhodné použít jiný nelineární model, ale ani výsledky vícerozměrné regrese nejsou zanedbatelné a mělo smysl regrese provést, jelikož byly odhaleny významné proměnné ovlivňující budoucí volatilitu. Dále bylo nutné ověřit předpoklady vícerozměrného regresního modelu. Z normálního p -grafu byla zjištěna mírně porušená podmínka normality. I z následného histogramu bylo vidět zešikmení dat. V závěru práce byly výsledky regresního modelu promítnuty do grafů. Cíl práce byl takto splněn. Klíčovým zjištěním bylo potvrzení vlivu sentimentu ve zprávách na volatilitu akcií s třídenním zpožděním. Další práce může být věnována delšímu časovému období, analýze více firem a použití dalších slovníků.

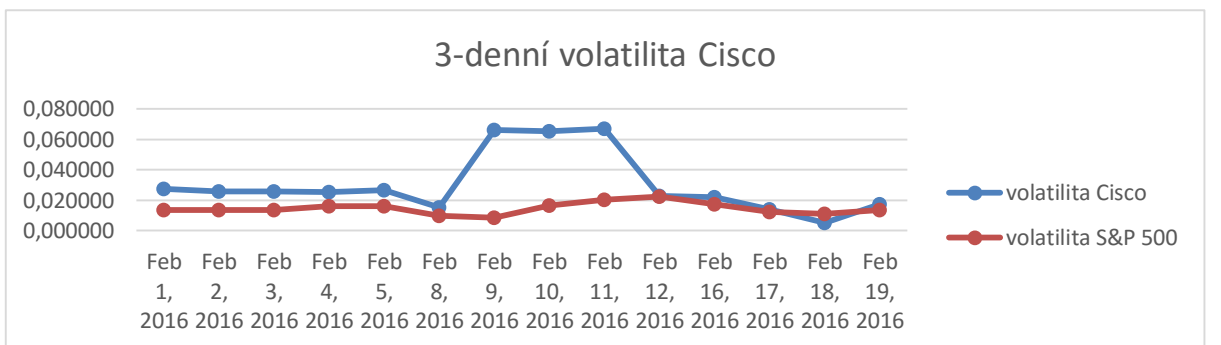
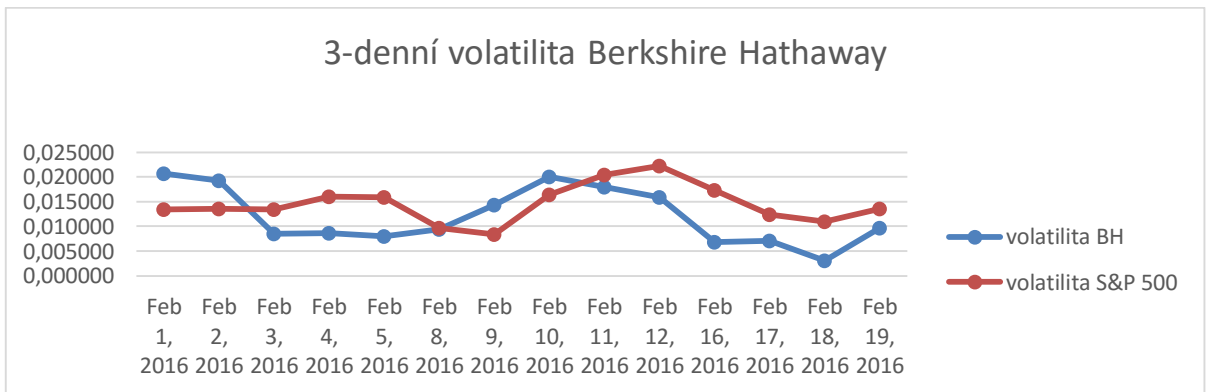
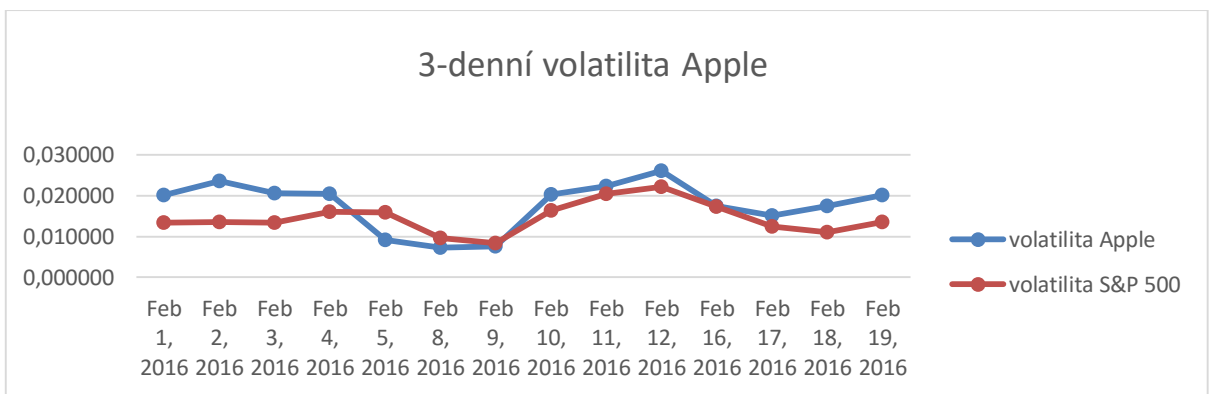
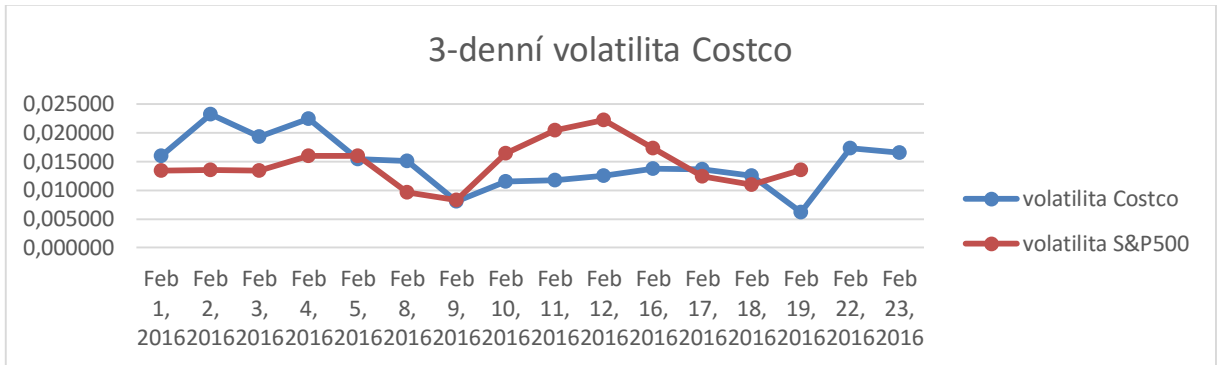
SEZNAM POUŽITÉ LITERATURY

- [1] EPSTEIN, Ira a David A. S. GARFIELD. *The psychology of smart investing: meeting the 6 mental challenges*. New York, N.Y.: J. Wiley, c1992. ISBN 047155071X.
- [2] GLADIŠ, Daniel. *Naučte se investovat*. Praha: Grada, 2004. Finanční trhy a instituce. ISBN 80-247-0709-8.
- [3] HINDLS, Richard, Stanislava HRONOVÁ a Jan SEGER. *Statistika pro ekonomy*. 2. vyd. Praha: Professional Publishing, 2002. ISBN 80-86419-30-4.
- [4] HNILICA, Jiří a Jiří FOTR. *Aplikovaná analýza rizika ve finančním managementu a investičním rozhodování*. Praha: Grada, 2009. Expert (Grada). ISBN 978-80-247-2560-4.
- [5] Jak na ukazatel beta? *Klubinvestoru* [online]. Flow media, 2014 [cit. 2016-06-27]. Dostupné z: <http://www.klubinvestoru.com/cs/article/2289-jak-na-ukazatel-beta>
- [6] JÍLEK, Josef. *Kapitálový a derivátový trh*. Praha: Bankovní institut, 1998. ISBN 80-7265-006-8.
- [7] KEARNEY, C., LIU, S. Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis*. 2014, roč. 23, č. 33, s. 170-178. ISSN 1057-5219
- [8] KOHOUT, Pavel. *Investiční strategie pro třetí tisíciletí*. 7., aktualiz. a přeprac. vyd. Praha: Grada, 2013. Finance (Grada). ISBN 978-80-247-5064-4.
- [9] KOTHARI, S.P., Xu LI a James E. SHORT. The Effect of Disclosures by Management, Analysts, and Business Press on Cost of Capital, Return Volatility, and Analyst Forecasts: A Study Using Content Analysis. *The Accounting Review*. 2009, roč. 84, č. 5, s. 1639-1670. DOI: 0.2308/accr.2009.84.5.1639.
- [10] LIŠKA, Václav a Jan GAZDA. *Kapitálové trhy a kolektivní investování*. Praha: Professional Publishing, 2004. ISBN 80-86419-63-0.
- [11] LIU, Bing a rafael SAX. *Sentiment analysis and opinion mining*. 2nd ed. San Rafael: Morgan, c2012, xiv, 165 s. Synthesis lectures on human language technologies, 16. ISBN 978-1-60845-884-4.
- [12] LOUGHRAN, T., MCDONALD, B. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*. 2011, roč. 66, č. 1, s. 35-65. ISSN 1540-6261.
- [13] MUSÍLEK, Petr. *Trhy cenných papírů*. 2., aktualiz a rozš. vyd. Praha: Ekopress, 2011. ISBN 978-80-86929-70-5.

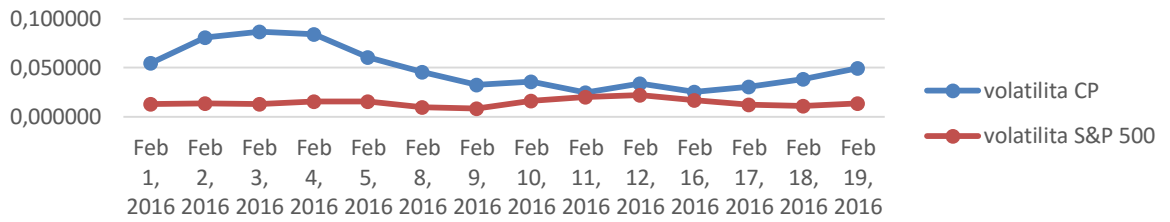
- [14] NEFTCI, Salih N. *Principles of financial engineering*. 2nd ed. Amsterdam: Elsevier, c2008, xiii, 679 s. Academic press advanced finance series. ISBN 978-012-3735-744.
- [15] On Social Sentiment and Sentiment Analysis. *Brnrđ: Everything digital* [online]. 2013 [cit. 2016-06-27]. Dostupné z: <http://brnrđ.me/social-sentiment-sentiment-analysis/>
- [16] PAVLÁT, Vladislav. *Kapitálové trhy*. Praha: Professional Publishing, 2003. ISBN 80-86419-33-9.
- [17] PLUMMER, Tony. *Prognóza finančních trhů: psychologie úspěšného investování*. Brno: Computer Press, 2008. ISBN 978-80-251-1592-3.
- [18] POON, Ser-Huang a Clive W. J. GRANGER. Forecasting Volatility in Financial Markets: A Review. *Journal of Economic Literature*. 2003, roč. 41, č. 2, s. 478-539.
- [19] RAMBOCAS, Meena a João GAMA. *Marketing research: The role of sentiment analysis* [online]. Universidade do Porto, č. 489. [cit. 2016-06-27]. ISSN 0870-8541. Dostupné z: <http://wps.fep.up.pt/wps/wp489.pdf>
- [20] REJNUŠ, Oldřich. *Finanční trhy*. 3., rozš. vyd. Ostrava: Key Publishing, 2011. Ekonomie (Key Publishing). ISBN 978-80-7418-128-3.
- [21] VESELÁ, Jitka. *Investování na kapitálových trzích*. 2., aktualiz. vyd. Praha: Wolters Kluwer Česká republika, 2011. ISBN 978-80-7357-647-9.
- [22] Volatilita. *Cz.saxobank* [online]. Praha, 2014 [cit. 2016-06-27]. Dostupné z: <http://cz.saxobank.com/support/slovník-pojmu/volatilita>
- [23] ZMEŠKAL, Zdeněk, Dana DLUHOŠOVÁ a Tomáš TICHÝ. *Finanční modely: koncepty, metody, aplikace*. 3., přeprac. a rozš. vyd. Praha: Ekopress, 2013. ISBN 978-80-86929-91-0.

PŘÍLOHY

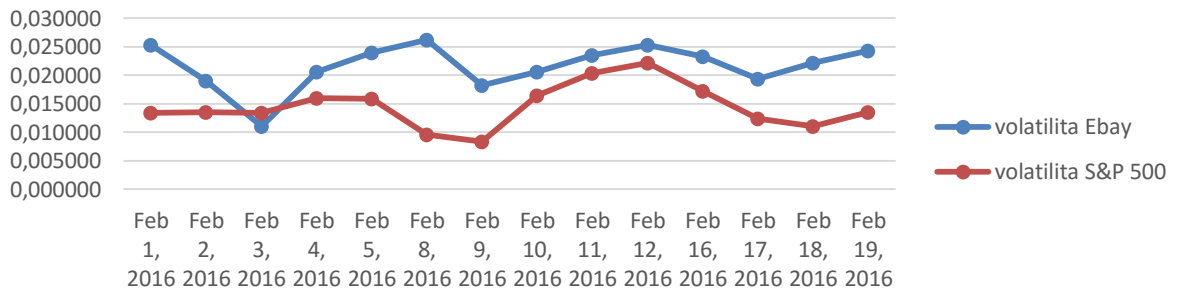
Příloha A – Porovnání volatilit firm s indexem S&P 500



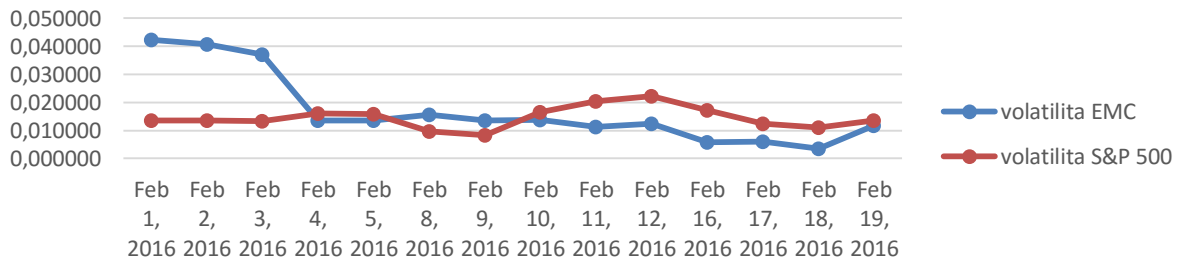
3-denní volatilita ConocoPhillips



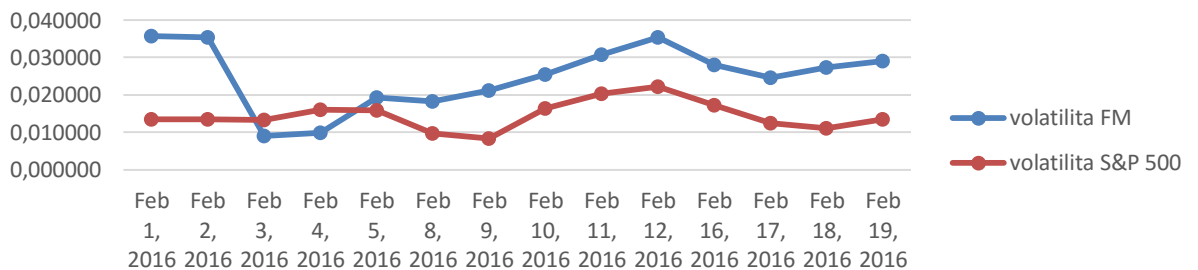
3-denní volatilita Ebay



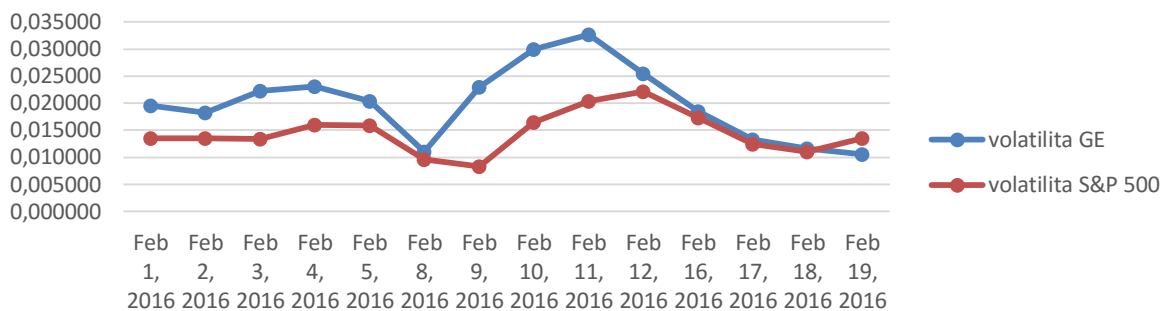
3-denní volatilita Exxon Mobil Corporation



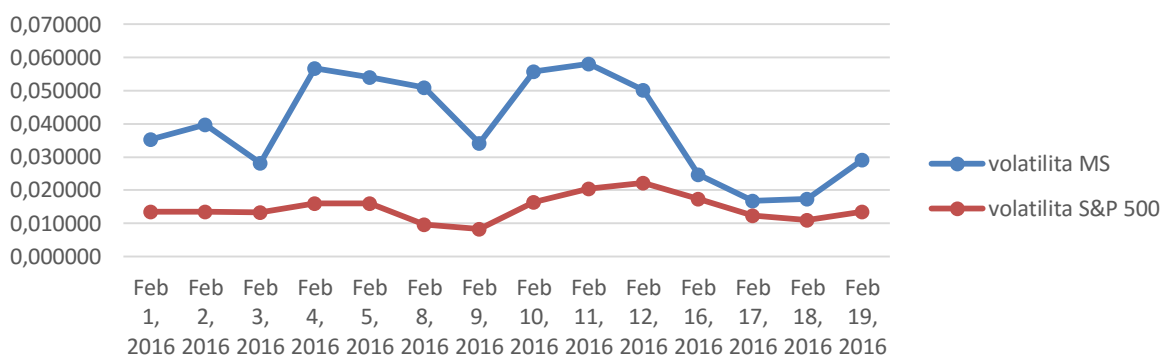
3-denní volatilita Ford Motor Co.



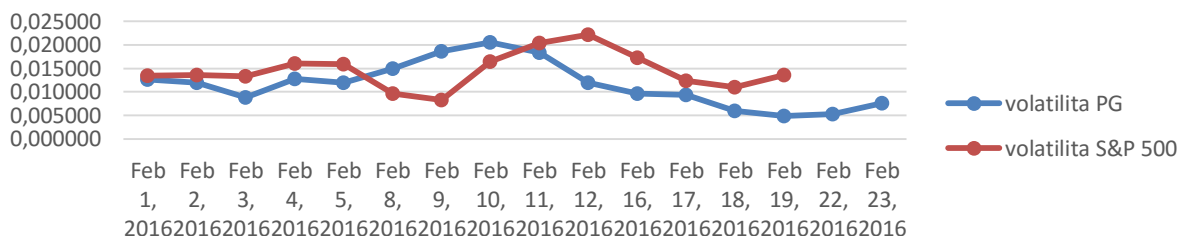
3-denní volatilita General Electric Company



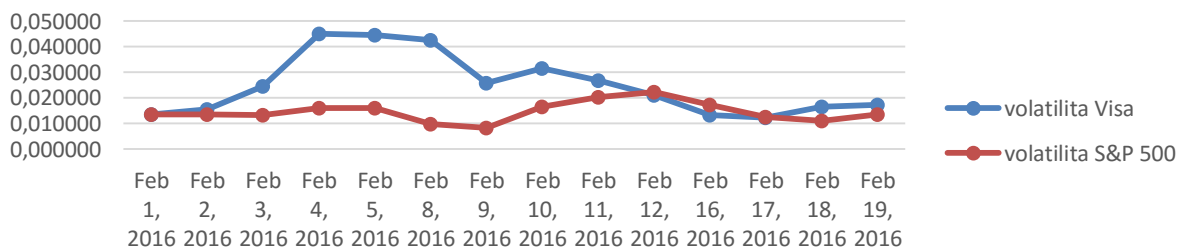
3-denní volatilita Morgan Stanley

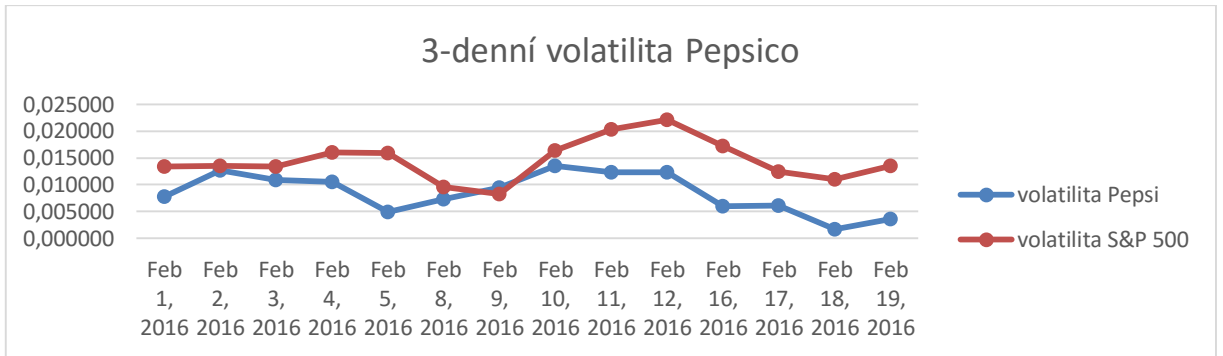
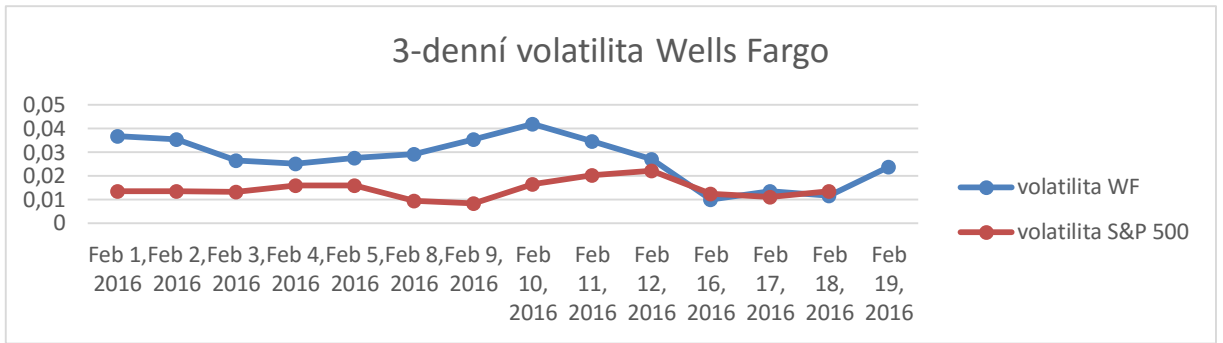


3-denní volatilita The Procter & Gamble Company

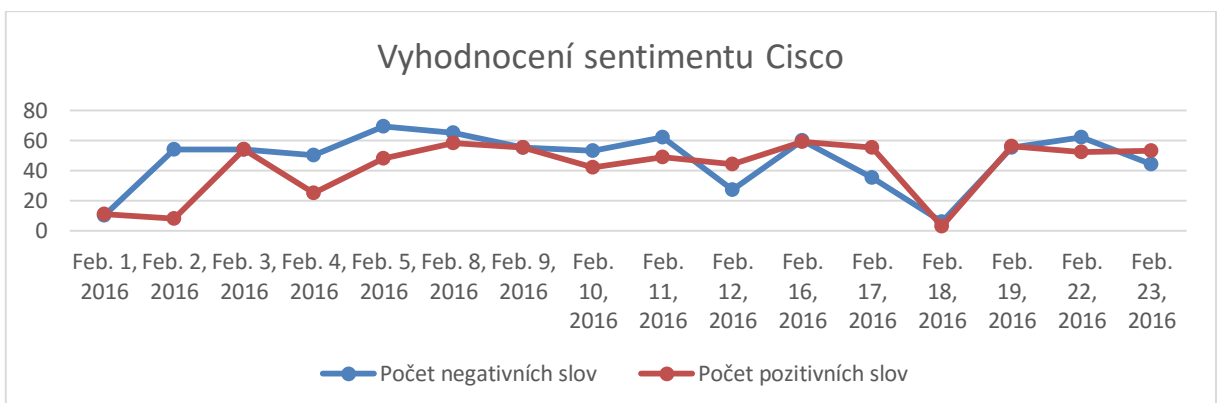
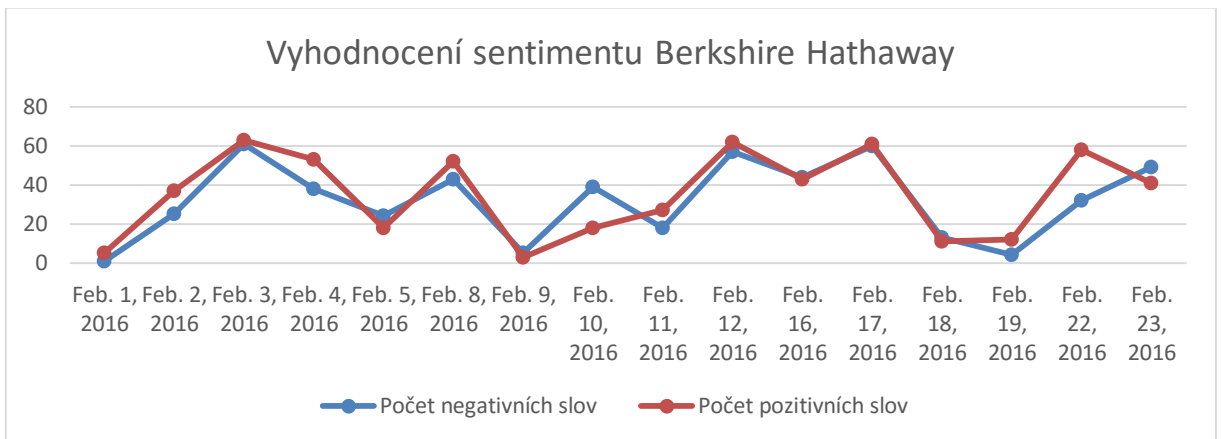
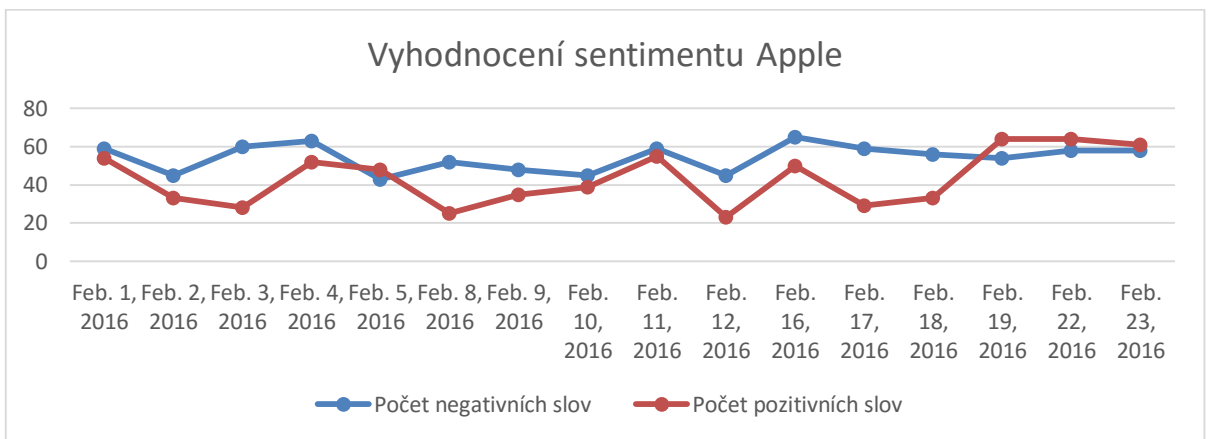
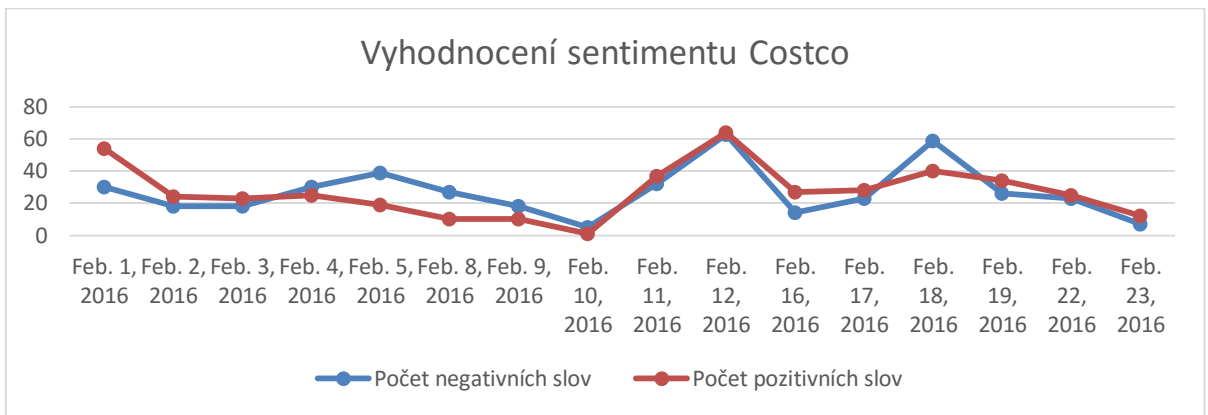


3-denní volatilita Visa

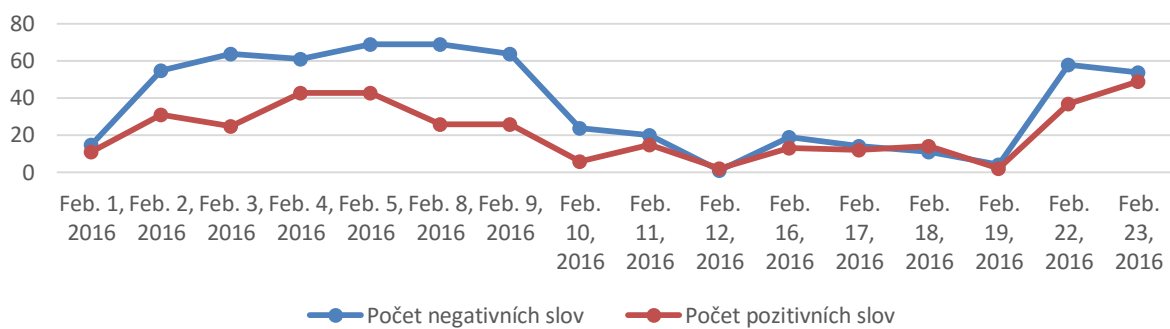




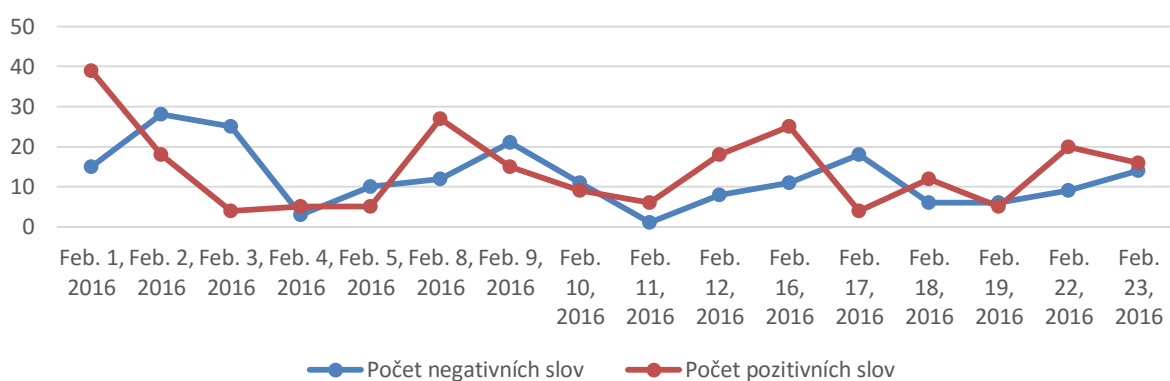
Příloha B – Vyhodnocení sentimentu firem



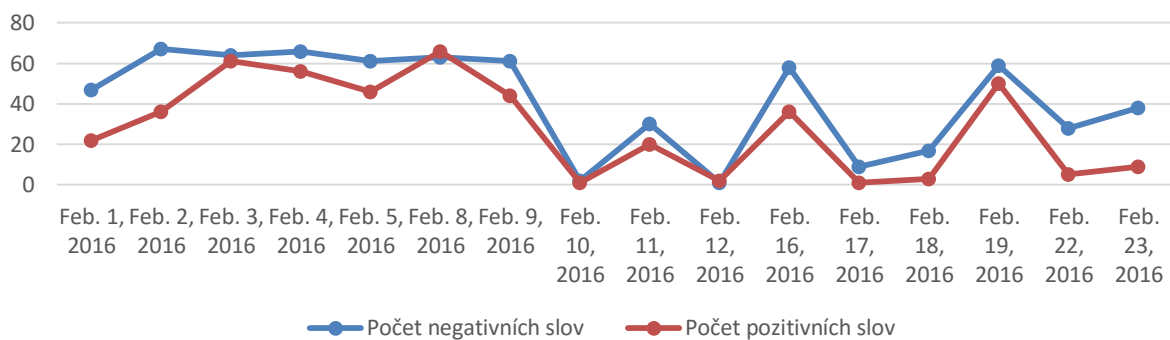
Vyhodnocení sentimentu Conoco Phillips



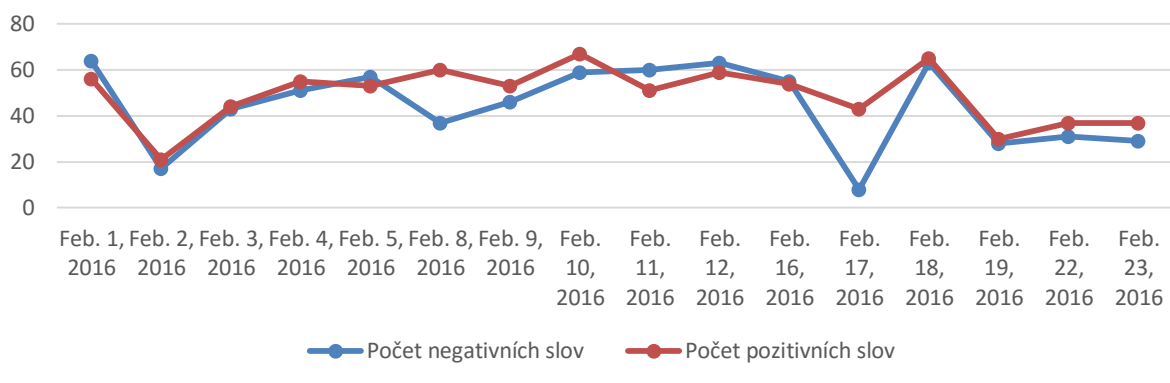
Vyhodnocení sentimentu Ebay



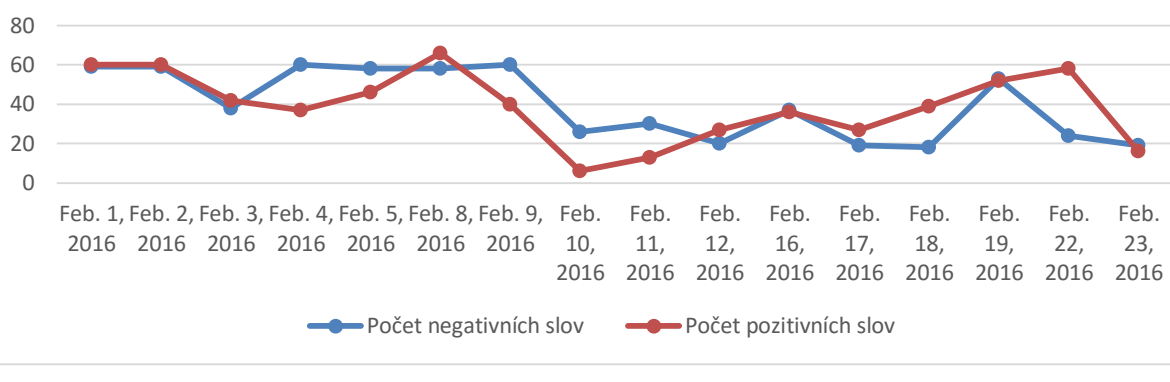
Vyhodnocení sentimentu Exxon Mobile Corporation



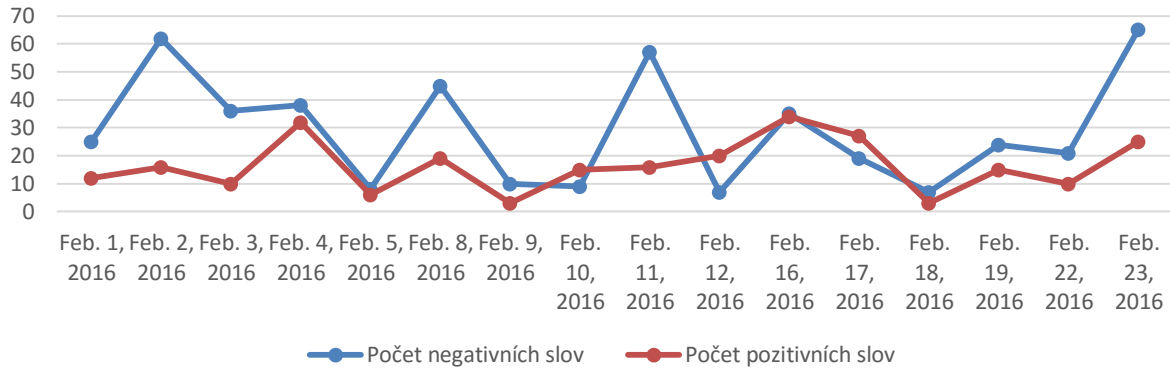
Vyhodnocení sentimentu Ford Motor



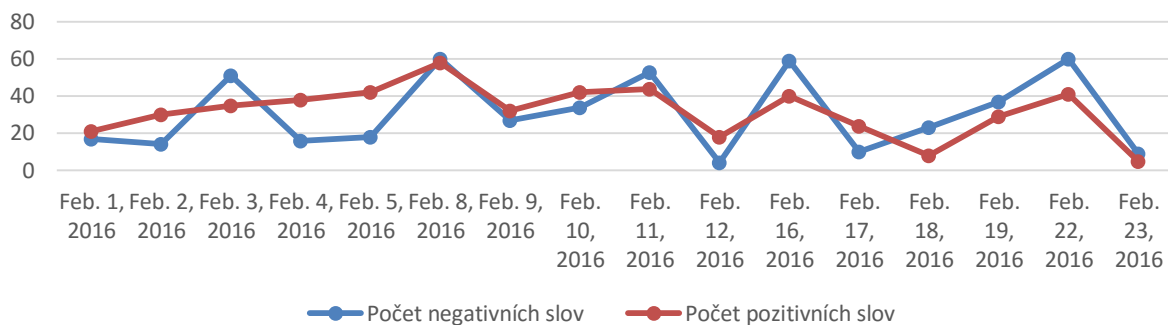
Vyhodnocení sentimentu General Electric



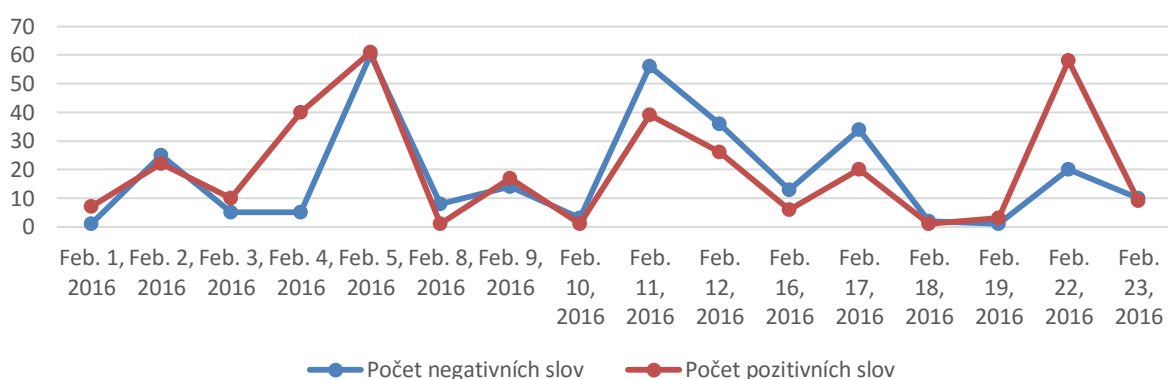
Vyhodnocení sentimentu Morgan Stanley



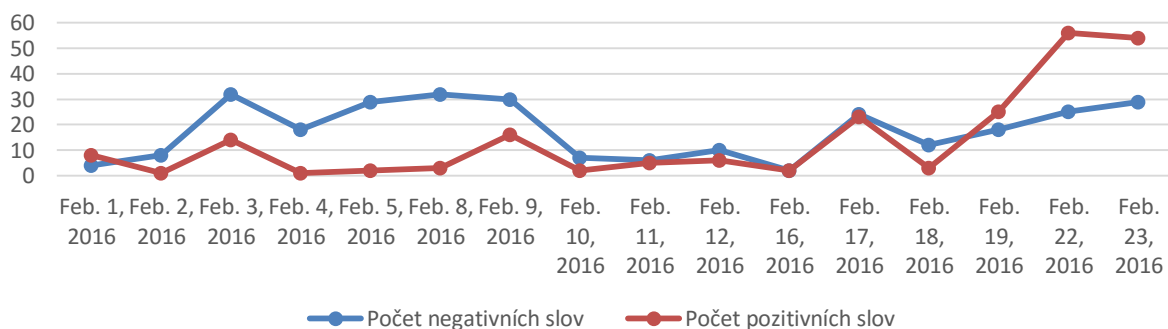
Vyhodnocení sentimentu The Protect & Gamble Company

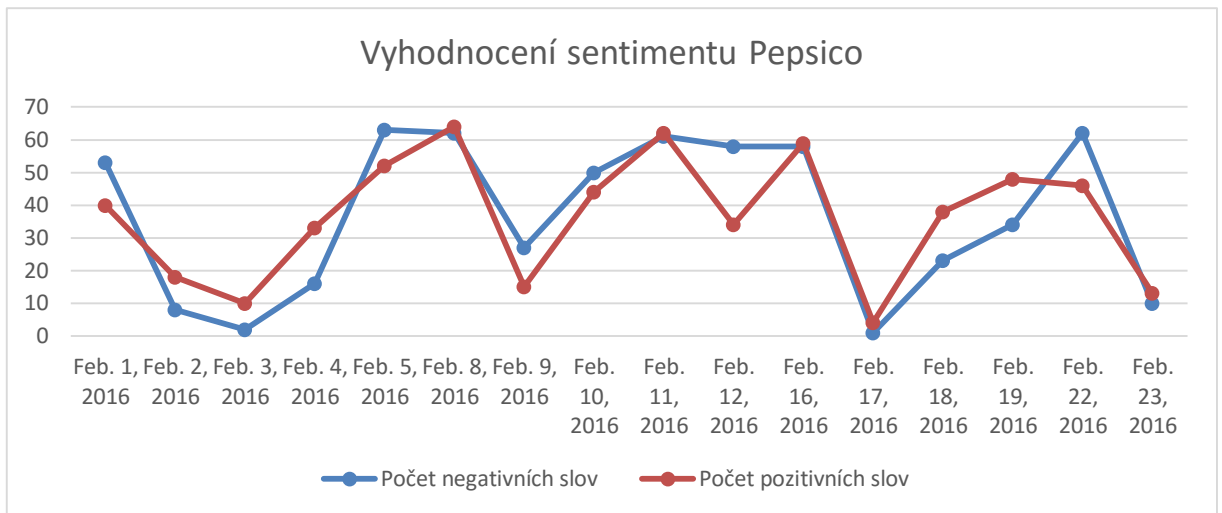


Vyhodnocení sentimentu Visa

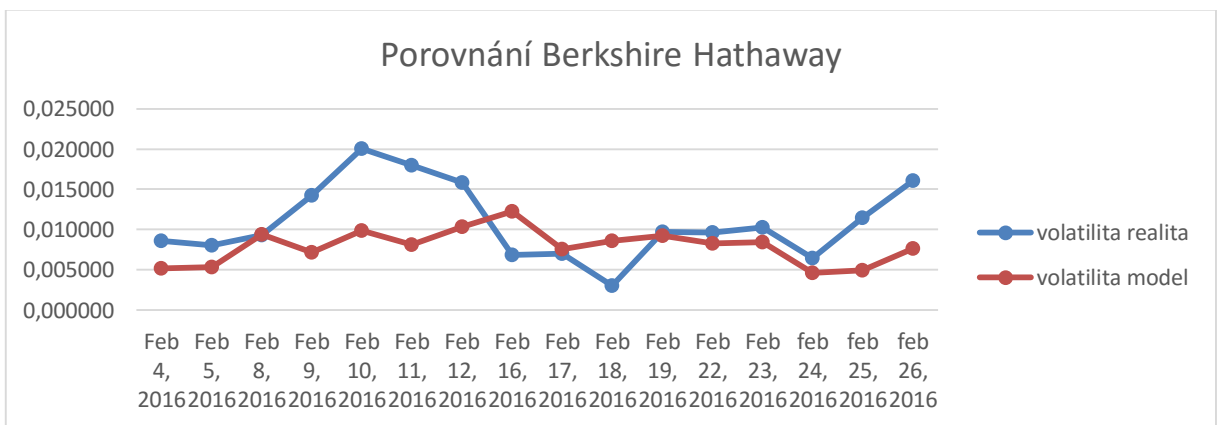
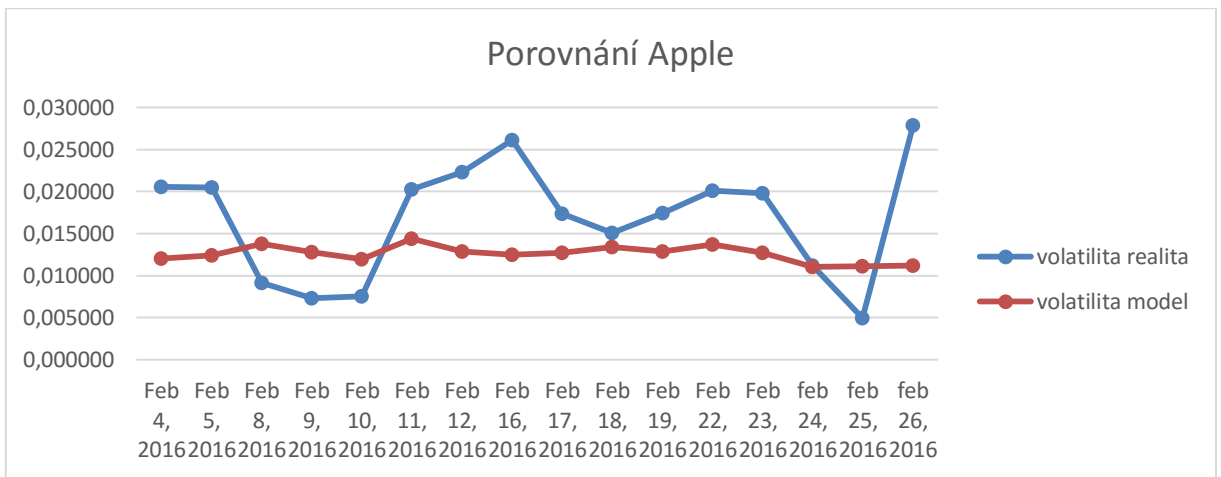
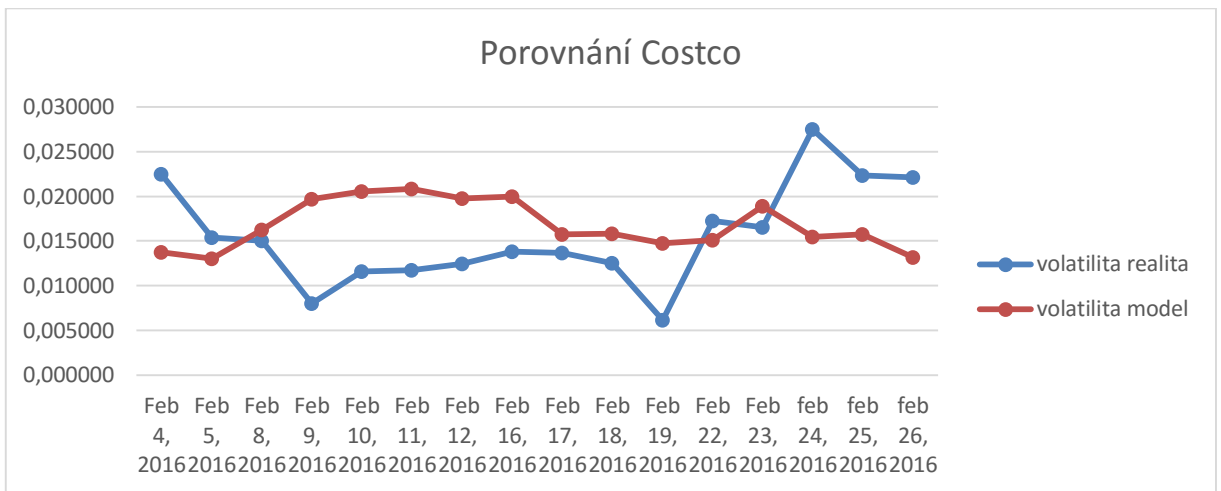


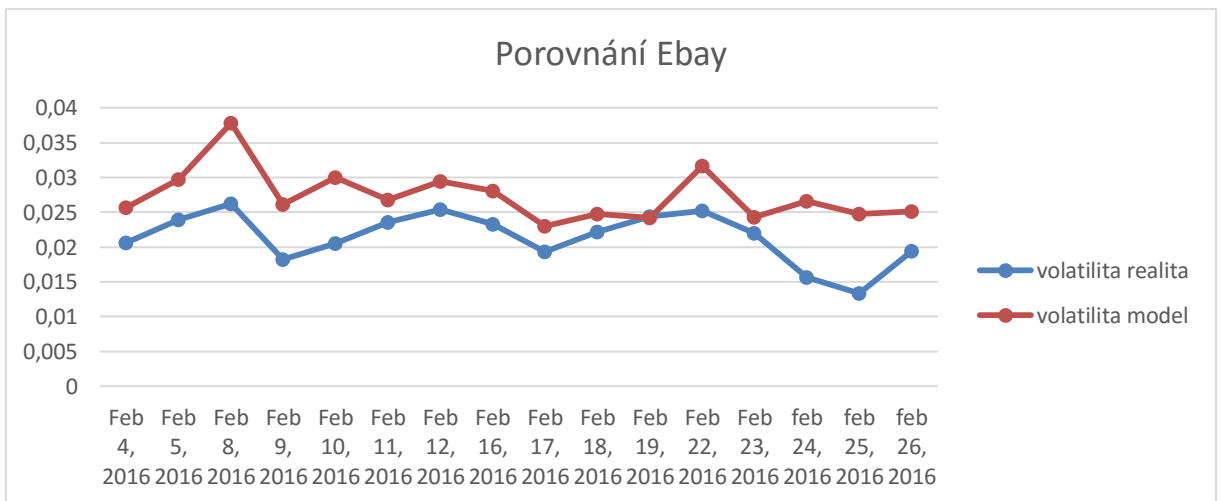
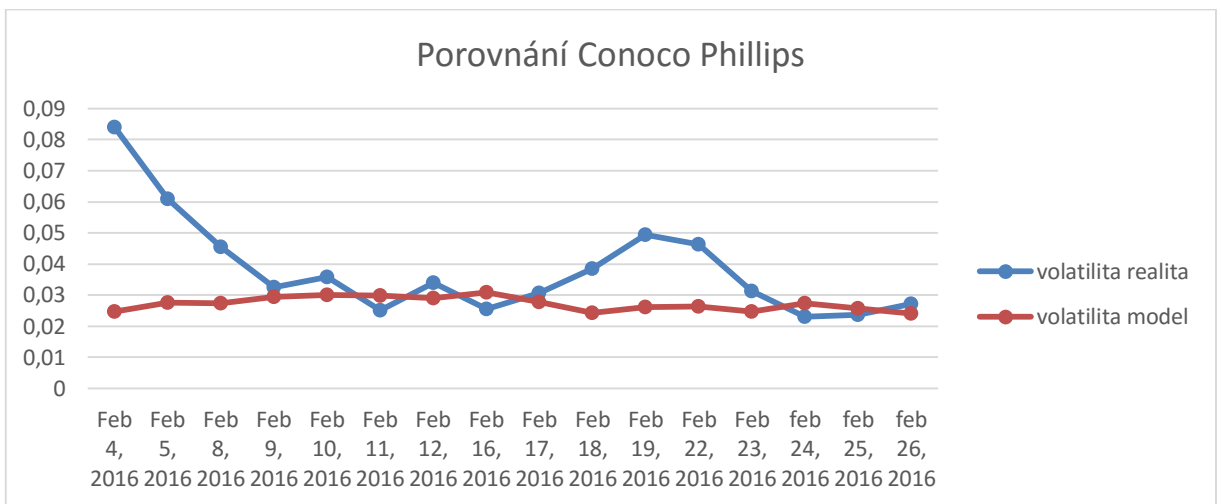
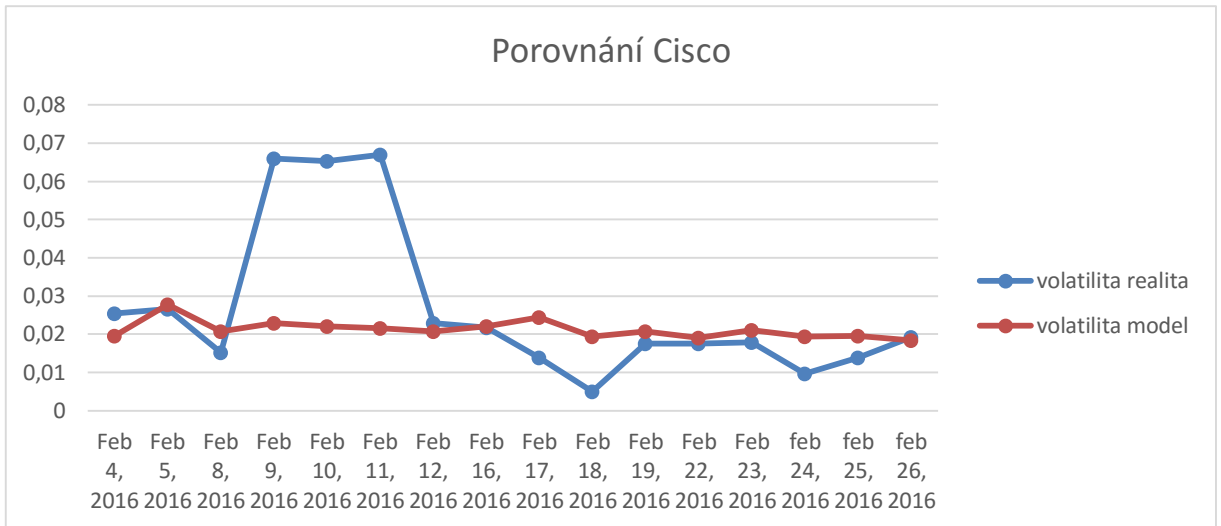
Vyhodnocení sentimentu Wells Fargo

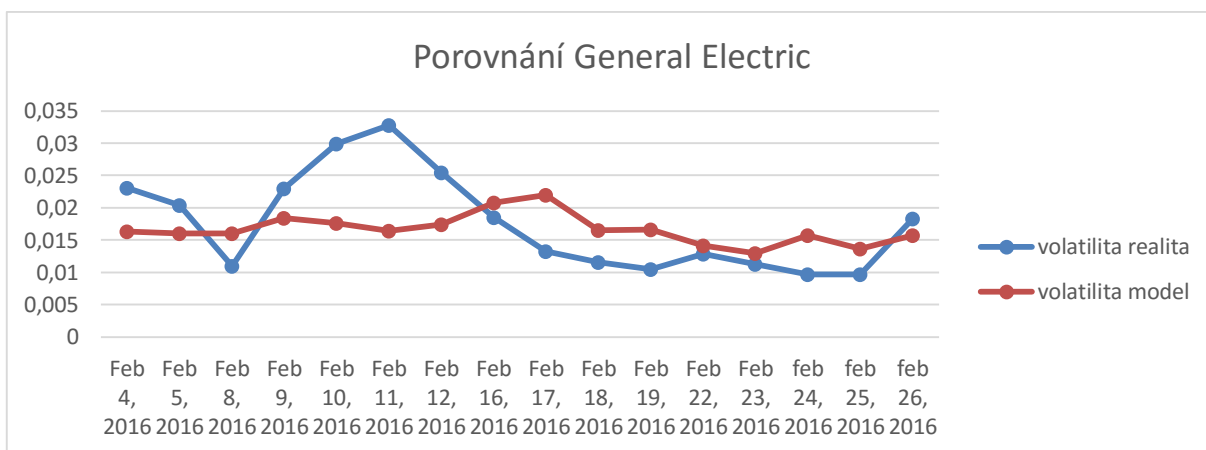
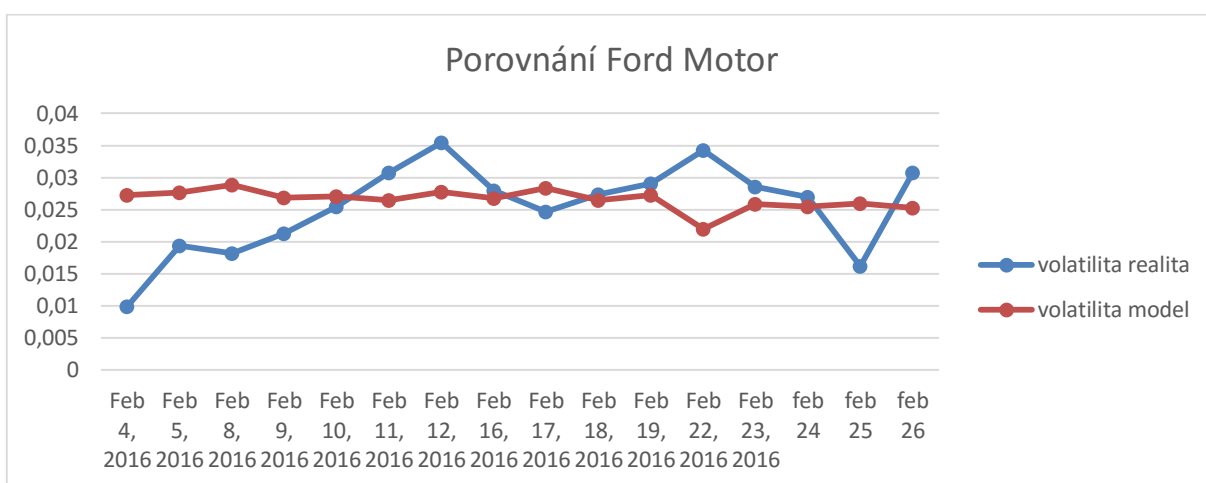
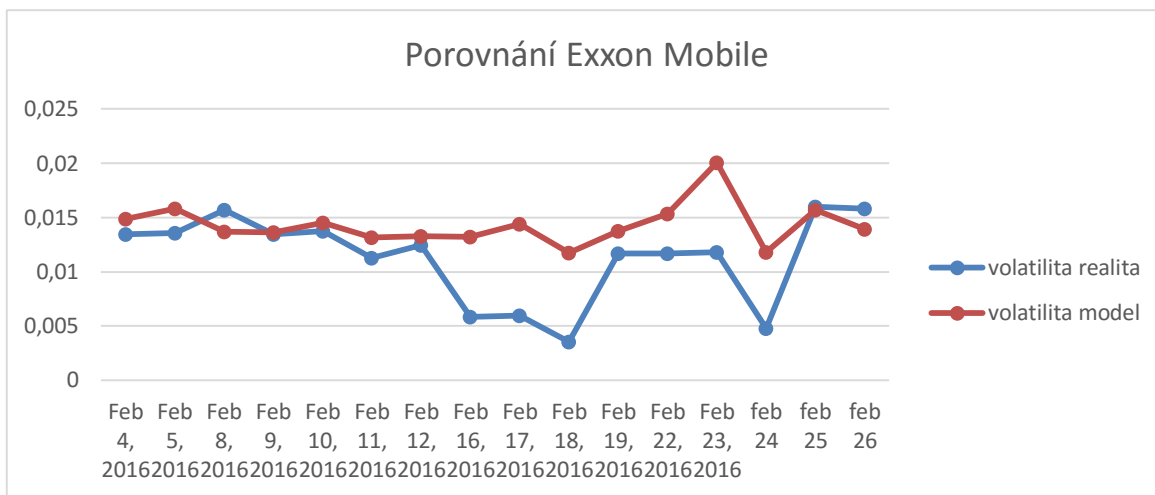


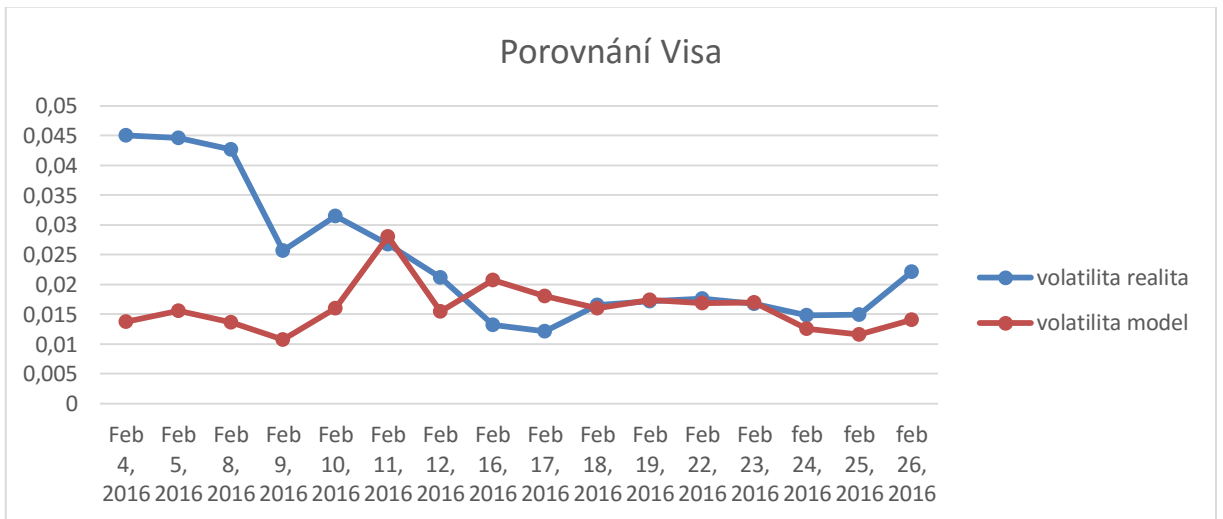
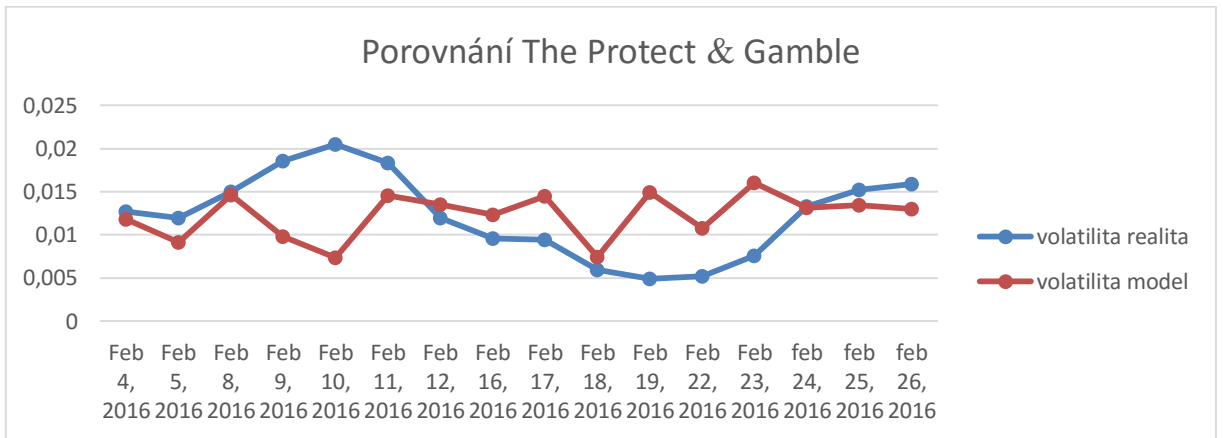
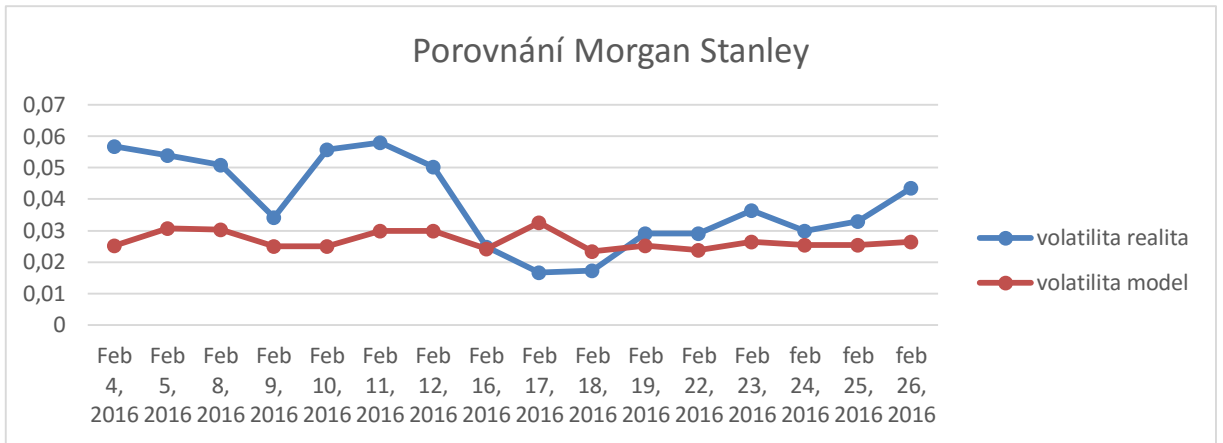


Příloha C – Vizualizace dosažených výsledků

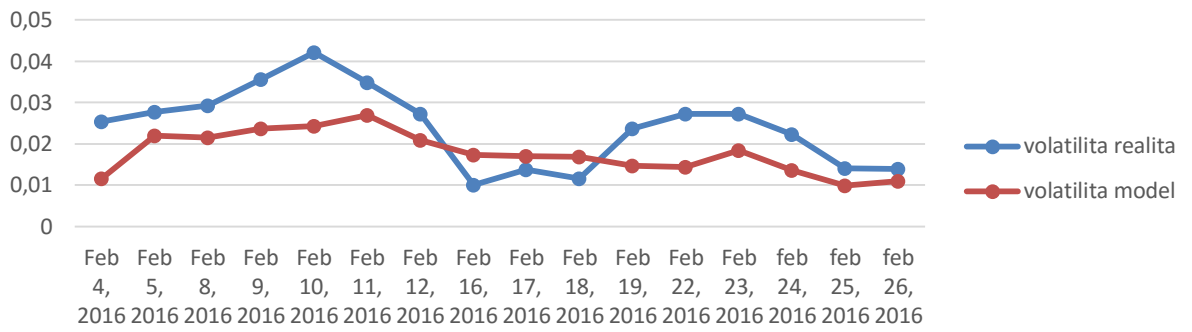








Porovnání Wells Fargo



Porovnání Pepsico

