

ERASMUS UNIVERSITY ROTTERDAM

ERASMUS SCHOOL OF ECONOMICS

BSc Economie en Bedrijfseconomie

Major Financial Economics

## Elon Musk's 'controversiële' Twitter-gebruik: de invloed op de aandelen van Tesla



Elon Musk   
@elonmusk



Tesla stock price is too high imo

5:11 p.m. · 1 mei 2020 · [Twitter for iPhone](#)

37,9K Retweets   203,2K Vind-ik-leuks



Bron: (@ElonMusk). 1 Mei 2020, 17:11. Tweet.

**Auteur:** Suzanne van den Brink

**Studentnummer:** 480994

**Scriptie begeleider:** Dr. Jan Lemmen

**Tweede lezer:** Amar Soebhag

## Samenvatting

In dit onderzoek is gekeken naar het effect van het Twitter-gebruik van Elon Musk op de aandelen van Tesla. Door middel van een Twitter-sentimentanalyse, verschillende lineaire regressieanalyses met Newey-West standaardfout en een Granger causality test, wordt het effect onderzocht. Uit dit onderzoek wordt geconcludeerd dat het gemiddeld positieve sentiment van de tweets kan leiden tot een positief effect op de aandelenkoers van Tesla. Ook is er een positief effect tussen de hoeveelheid tweets en het handelsvolume van Tesla gevonden, maar tussen de hoeveelheid tweets en bid-ask spread is geen significant negatief effect waargenomen. Het is dus onzeker of tweets een effect hebben op de liquiditeit van Tesla-aandelen. Er is een causaal verband gevonden tussen het aantal Tweets en handelsvolume. Dit betekent dat de hoeveelheid Tweets een voorspeller is van het handelsvolume. De conclusie van het onderzoek is dat het Twitter-gebruik van Elon Musk geen bewezen effect heeft op de volatiliteit van Tesla-aandelen en een onzeker positieve invloed heeft op de liquiditeit van het Tesla-aandeel. De grootste tekortkoming van het onderzoek is het gebrek aan data in 2018.

*Trefwoorden: Elon Musk, Tesla, Twitter, sentiment, liquiditeit*

## Inhoud

|  |    |
|--|----|
| Samenvatting.....  | 2  |
| Hoofdstuk 1: Inleiding .....                             | 5  |
| Hoofdstuk 2: Literatuuroverzicht .....                   | 7  |
| 2.1 Twitter .....  | 7  |
| 2.2 Twitter en CEO's .....                               | 7  |
| 2.3 Twitter en de aandelenmarkt.....                     | 8  |
| 2.3.1 Twitter-sentiment en de aandelenmarkt.....         | 8  |
| 2.4 Hypotheses.....                                      | 10 |
| 2.4.1 Hypothese 1.....                                   | 10 |
| 2.4.2 Hypothese 2.....                                   | 10 |
| 2.4.3 Hypothese 3.....                                   | 10 |
| Hoofdstuk: 3 Methodologie .....                          | 11 |
| 3.1 Hypothese 1.....                                     | 11 |
| 3.2 Hypothese 2.....                                     | 11 |
| 3.3 Hypothese 3.....                                     | 12 |
| Hoofdstuk: 4 Data.....                                   | 13 |
| 4.1 Begrippen .....                                      | 13 |
| 4.2 Twitteraccount Elon Musk.....                        | 13 |
| 4.2.1 Twitter-gebruik Elon Musk .....                    | 14 |
| 4.3 Variabelen .....                                     | 19 |
| 4.3.1 Stationariteit.....                                | 20 |
| Hoofdstuk 5: Resultaten.....                             | 22 |
| 5.1 Hypothese 1.....                                     | 22 |
| 5.2 Hypothese 2.....                                     | 23 |
| 5.2.1 Handelsvolume .....                                | 23 |
| 5.2.2 Bid-ask spread .....                               | 24 |
| 5.2.3 Liquiditeit.....                                   | 25 |
| 5.3 Hypothese 3.....                                     | 26 |
| Hoofdstuk 6: Conclusie, Discussie en Aanbevelingen ..... | 27 |
| 6.1 Conclusie .....                                      | 27 |
| 6.2 Discussie en aanbevelingen.....                      | 28 |
| Bibliografie .....                                       | 29 |
| Literatuurbronnen.....                                   | 29 |
| Internetbronnen.....                                     | 30 |
| Appendix.....  | 32 |

## Effect van Elon Musk's Twitter-gebruik

|   |    |
|---|----|
| Appendix I – STATA Do-File .....                      | 32 |
| Appendix II – STATA output.....                       | 34 |
| Appendix III – Python codes .....                     | 40 |
| Code 1: Tweets genereren .....                        | 40 |
| Code 2: VADER sentiment .....                         | 41 |
| Code 3: Word cloud.....                               | 42 |
| Appendix IV: Relevante python modules + versies ..... | 43 |

## Hoofdstuk 1: Inleiding

Op 7 augustus 2018 verdiende Elon Musk, CEO van Tesla Inc., \$1,4 miljard door slechts negen woorden op Twitter te plaatsen: “*Am considering taking Tesla private at \$420. Funding secured.*” (@ElonMusk). 7 Augustus 2018, 18:48. Tweet.). Hierdoor schoot de aandelenprijs van Tesla de lucht in. Uren na de tweet, een kort berichtje van maximaal 140 karakters, werd de handel in Tesla-aandelen voor 90 minuten opgeschort en werd op de officiële website van Tesla een uitleg geplaatst over de privatisering. In de uitleg werd omschreven hoe privatisering leidt tot betere prestaties van het bedrijf (Musk, 2018). Aan het einde van de dag was de koers gestegen met 11% (Money, 2018) en het handelsvolume steeg tot zeven miljoen; het hoogste handelsvolume sinds 2014 (Levy, 2018). Een maand later plaatste Elon Musk een tweet waarin hij afzag van de plannen om Tesla te privatiseren. Door sommigen wordt verondersteld dat de tweet bedoeld was als grap en een referentie was naar 4/20, een ‘feestdag’ voor cannabisgebruikers.

De *Securities and Exchange Commission*, de Amerikaanse instantie die beurzen reguleert, reageerde op deze tweet met een aanklacht tegen Musk wegens fraude door koersmanipulatie (SEC, Elon Musk Charged With Securities Fraud for Misleading Tweets, 2018). Tesla heeft in 2013 namelijk bij de SEC vastgelegd dat @ElonMusk bedoeld is als platform voor officieel nieuws over Tesla (Simpson, 2018), de \$420-tweet werd dus opgevat als betrouwbaar nieuws en niet als grap. De aanklacht van SEC leidde tot een schikking, waarbij een boete van \$40 miljoen werd geëist en de aftreding van Musk als chairman voor de komende drie jaar (SEC, 2018). De helft van de boete werd betaald door Musk zelf, de andere helft werd betaald door Tesla. De schikking leidde tot een koersdaling van 14% voor Tesla (Kelleher, 2018). Een jaar later ging Musk wederom in overtreding. Musk plaatste namelijk wederom een misleidende tweet, die leidde tot koersstijging. In deze tweet verkondigde hij dat Tesla om en nabij 500.000 auto's zou produceren in 2019, terwijl het begrote aantal tussen de 360.000 en 400.000 auto's was. Investeerders werden dus geconfronteerd met een te positieve valse voorspelling. Uren later kwam Musk terug op deze voorspelling en verwijderde hij de tweet. Na deze tweede misleidende tweet kwam Musk tot een overeenkomst met de SEC: voortaan zou een advocaat de tweets over Tesla's financiën, verkopen en afleveringen goedkeuren voordat deze openbaar gemaakt worden (O'Keane, 2019).

Tegenwoordig is Twitter de beste manier voor CEO's om in contact te komen met klanten en aandeelhouders (Malholtra & Malholtra, 2016). Musk maakt hier gebruik van door af en toe nieuwe informatie over Tesla en SpaceX te delen op Twitter. Doordat Musk dit soms exclusief deelt op Twitter zorgen zijn tweets voor speculaties en media-aandacht. Volgens Malhotra en Malhotra zou het gebruik van Twitter door CEO's enkele voordelen hebben. Echter, door de misleidende tweets op het persoonlijke account van Musk hebben Tesla en Musk ook nadelen ondervonden van het Twitter-

## Effect van Elon Musk's Twitter-gebruik

gebruik van Musk. Tegelijkertijd zorgen zijn tweets ook voor koersstijgingen en publiciteit voor Tesla. In dit onderzoek wordt het effect tussen het Twitter-gebruik van Elon Musk en het effect op de aandelen van Tesla onderzocht. Dit leidt tot de volgende onderzoeksvraag:

*Leidt het Twitter-gedrag van Elon Musk tot een positief of negatief effect op de prestatie van Tesla Inc. aandelen?*

Om deze vraag te beantwoorden zal allereerst een literatuuronderzoek dat focust op eerder empirisch onderzoek naar Twitter en de aandelenmarkt en Twitter-sentiment uitgevoerd. Hierna wordt er een set van drie hypothesen getest, dit met behulp van regressie-analyses en een Granger causality test. Na het testen van de hypothesen kan een conclusie worden getrokken over de invloed van het Twitter-gebruik van Elon Musk op de aandelen van Tesla. Het onderzoek wordt afgesloten met discussie en aanbevelingen voor toekomstig onderzoek.

## Hoofdstuk 2: Literatuuroverzicht

### 2.1 Twitter

Twitter is een website waarop gebruikers sinds 2006 berichten kunnen plaatsen. Het platform wordt ook wel een microblog genoemd, omdat gebruikers maximaal 140 karakters kunnen gebruiken per bericht. Via Twitter worden op elk moment van de dag nieuws en belangrijke gebeurtenissen gedeeld via Twitter (De Maio, Fenza, Loia, & Parente, 2016). Maandelijks worden er door 330 miljoen gebruikers gebruik gemaakt van de services die Twitter biedt. Van deze maandelijkse gebruikers is ongeveer de helft (152 miljoen) dagelijks actief. Per dag worden er gemiddeld 500 miljoen tweets verstuurd (Aslam, 2020). Tweets zijn zichtbaar voor gebruikers en niet-gebruikers van het platform wanneer een account op openbaar staat. Een gebruiker kan er ook voor kiezen om het account op privé te zetten; dan is er toestemming nodig van de eigenaar om het account in te mogen zien en te kunnen volgen. Het account van Elon Musk is openbaar en telt op dit moment bijna 40 miljoen volgers (Twitter, 2020).

Bij de oprichting van Twitter, in 2006, was het platform ontworpen om berichten te delen via SMS, vandaar het limiet van 140 karakters per bericht. Door de omschakeling van SMS naar een desktopwebsite en, uiteindelijk, een mobiele applicatie steeg het gebruikersgemak van Twitter en nam het aantal gebruikers toe (Liu, Kliman-Silver, & Mislove, 2014). In 2007 werden *hashtags*, aangeduid met een #, geïntroduceerd. Hiermee kan men belangrijke onderwerpen indexeren en worden de meest gebruikte onderwerpen weergegeven bij *trending* (Twitter, 2020). Sinds 2009 zijn de hoeveelheid *mentions* (antwoorden op tweets) toegenomen: tegenwoordig zijn mentions goed voor 50% van de totale tweets. Dit zorgde voor een meer interactief karakter tussen de gebruikers van Twitter.

### 2.2 Twitter en CEO's

In 2013 gaf de SEC groen licht aan bedrijven om (belangrijke) informatie te delen op sociale media, onder voorwaarde dat investeerders op de hoogte gebracht zijn op welk sociale media platform deze informatie gedeeld zal worden (SEC, SEC Says Social Media OK for Company Announcements if Investors Are Alerted, 2013). Hier maakt Elon Musk gebruik van door op zijn persoonlijke Twitter-account regelmatig nieuws te delen over Tesla. Musk is hierin één van de weinige CEO's, maar het aantal CEO's dat nieuws deelt via Twitter groeit. In 2015 gebruikte nog geen 9% van de *Fortune 500* CEO's Twitter (Malholtra & Malholtra, 2016), dit is in 2019 gestegen naar 14% (Clinton, 2019).

CEO's kunnen via Twitter een positieve invloed uitoefenen. Zo kunnen zij slecht nieuws over een bedrijf verzachten via een persoonlijke tweet, waardoor het vertrouwen van investeerders wordt behouden (Elliott, Grant, & Hodge, 2018). Daarnaast is er een aantoonbare relatie tussen verkochte aandelen en investeerders sentiment: individuele investeerders kopen vaak aandelen die de

## Effect van Elon Musk's Twitter-gebruik

aandacht trekken (Barber & Odean, 2008). Aandelen die, bijvoorbeeld, in het nieuws zijn geweest worden hierdoor vaker gekocht door investeerders. Dit kan ook betekenen dat investeerders die vaak nieuws over een bedrijf op Twitter zien eerder geneigd zijn deze aandelen te kopen. Hetzelfde geldt voor 'controversiële' tweets die het nieuws bereiken, hierdoor stijgt de vraag van het aandeel wat leidt tot positieve prijsdruk.

### 2.3 Twitter en de aandelenmarkt

Officiële bekendmakingen van een bedrijf bereiken vaak maar een deel van de investeerders; dit leidt tot informatie-asymmetrie. Door deze bekendmakingen ook via Twitter te verspreiden, leidt dit tot tot een lagere bid-ask spread en een groter handelsvolume. Er is sprake van een verminderde informatie asymmetrie (Blankespoor, Miller, & Hal, 2014). Het handelsvolume, de bid-ask spread en de absolute prijsverandering van een aandeel zijn sterk gecorreleerd met de hoeveelheid tweets van en over het bedrijf. Wanneer, gedurende de dag, de prijs van een aandeel sterk stijgt of daalt, stijgt het aantal tweets over deze aandelen. Dit effect is bewezen in een onderzoek naar Apple Inc., het bedrijf waarover dagelijks het meest getweet wordt van alle S&P 500 bedrijven (Mao, Wang, Wei, & Liu, 2012). Er kan nog meer gedaan worden met het nieuws die tweets bevatten. Naast het effect van tweets op de aandelenprijs kunnen meningen uit tweets gebeurtenissen voorspellen, zoals abnormale winsten die zich voordoen rond bekendmakingen van kwartaalcijfers van een bedrijf (Bartov, Faurel, & Mohanram, 2018).

#### 2.3.1 Twitter-sentiment en de aandelenmarkt

Er zijn meerdere onderzoeken waaruit blijkt dat de aandelenmarkt voorspeld kan worden door middel van Twitter-data. Uit deze data kan ook Twitter sentiment worden gehaald: de emoties die een tweet bevat. In een onderzoek van Zhang, Fuehres en Gloor (2011) is gebruik gemaakt van twee stemmingen: hoop en angst. Er is geconcludeerd dat een negatieve (hoopvolle) stemming in tweets leidt tot een daling (stijging) van aandelenindexen (S&P500, Dow Jones en NASDAQ) (Zhang, Fuehres, & Gloor, 2011). Dit effect op de aandelenmarkt is een combinatie van het nieuws dat een tweet bevat en het sentiment van de gebruiker. Bollen, Mao en Zheng (2011) onderzochten dit effect ook met het algemene sentiment van Twitter en de Dow Jones. Ze onderscheiden hierbij zes stemmingen: kalmte, alertheid, zekerheid, vitaliteit, vriendelijkheid en blijdschap. Ook in dit onderzoek wordt een significant effect waargenomen tussen het sentiment en de Dow Jones index (Bollen, Mao, & Zheng, 2011). Het gevoel dat een tweet bevat is dus een voorspeller van de aandelenmarkt. Volgens Sul, Dennis & Yuan (2016) is dit een langzaam proces en wordt het sentiment langzaam opgenomen in de aandelenprijs. Het sentiment over een bepaald bedrijf van een Twitter-gebruiker leidt tot een significant effect op de aandelenprijs van één dag later, tien dagen



## Effect van Elon Musk's Twitter-gebruik

later en twintig dagen later. Een strategie gebaseerd op Twitter-sentiment zou kunnen leiden tot een jaarlijkse winst van 11% tot 15% (Sul, Dennis, & Yuan, 2016).

CEO's kunnen hun bedrijf naar een hoger niveau tillen en een positief sentiment achterlaten bij de investeerder (Malholtra & Malholtra, 2016). Wanneer investeerders geconfronteerd worden met Elon Musk's positieve (negatieve) stemming op Twitter, kan dit zorgen voor een positief (negatief) effect op de aandelenprijs. Uit eerdere genoemde tweets van Elon Musk kan geconcludeerd worden dat zijn tweets een effect hebben op de aandelenkoers van Tesla. Zo leidde de \$420-tweet tot een grote stijging van de aandelenkoers. Men beschouwt de inhoud van Elon Musk's tweets als 'nieuws' over Tesla. Tweets hebben dus invloed op investeerders; Twitter-accounts worden niet alleen gebruikt als nieuwsplatform, maar worden ook gebruikt als beïnvloeder/voorspeller van de aandelenmarkt.

## 2.4 Hypotheses

De onderzoeksvraag luidt als volgt: 'Leidt het Twitter-gedrag van Elon Musk tot een positief of negatief effect op de prestatie van Tesla Inc. aandelen?'. Deze vraag zal worden beantwoord door middel van een set van drie hypothesen.

### 2.4.1 Hypothese 1

Allereerst wordt gekeken naar het directe effect dat het Twitter-gebruik van Elon Musk heeft op de aandelenkoers van Tesla. Hierbij wordt er niet specifiek gekeken naar de inhoud van de tweets, maar naar de invloed van het algemene gebruik. Dit leidt tot de eerste hypothese:

*De hoeveelheid tweets van Elon Musk is positief gecorreleerd met de volatiliteit van Tesla-aandelen.*

### 2.4.2 Hypothese 2

Vervolgens wordt getest in welke mate investeerders beïnvloed worden door de tweets van Elon Musk. Dit wordt onderzocht met behulp van de liquiditeit van het aandeel. De liquiditeit wordt onderzocht aan de hand van het handelsvolume en de bid-ask spread van het Tesla-aandeel. De tweede hypothese luidt:

*De hoeveelheid tweets van Elon Musk is positief gecorreleerd met de liquiditeit van Tesla-aandelen.*

### 2.4.3 Hypothese 3

Gebaseerd op de zojuist genoemde hypothesen zal er getest worden of de tweets van Elon Musk niet alleen correleren (indien dit het geval is), maar ook of er een causaal verband is.

*De hoeveelheid tweets van Elon Musk bevat voorspellende waarde over de volatiliteit en liquiditeit.*

## Hoofdstuk: 3 Methodologie

### 3.1 Hypothese 1

Om te testen of de volatiliteit van Tesla-aandeel gecorreleerd is met het aantal tweets van Elon Musk, zal er een (robuuste) enkelvoudige regressieanalyse worden gemaakt. Deze zal de volgende vorm hebben:

$$\text{Lineaire regressieanalyse } \sigma_{e_i} = \alpha + \beta T + u$$

Wanneer er een significant effect gevonden wordt tussen de afhankelijke variabele ( $\sigma_{e_i}$ ) en de afhankelijke variabele ( $\beta$ ) wordt, onder de nulhypothese van geen effect,  $H_0$  verworpen.

$$H_0: \beta = 0$$

$$H_a: \beta \neq 0$$

De foutenterm wordt robuust gemaakt om te verzekeren dat 'outliers' geneutraliseerd worden en geen onverwachte uitkomsten geven. In de regressie is de volatiliteit van Tesla's aandelen de afhankelijke variabele, deze wordt berekend aan de hand van de volgende formules:

$$\text{Variantie } (\sigma^2) : (X - \mu)^2 / N$$

$$\text{Volatiliteit } (\sigma) : \sqrt{\sigma^2}$$

De onafhankelijke variabele is de hoeveelheid tweets van het Twitter-account @ElonMusk. Meerdere controlevariabelen worden getest op significantie: handelsvolume, bid-ask spread, Tesla's schuld ratio ( $D/E$ ), return Tesla, return van de NASDAQ-100 en marktvolatiliteit.

### 3.2 Hypothese 2

Om te testen of de liquiditeit van Tesla's aandelen gecorreleerd is met het aantal tweets van Elon Musk zullen er twee variabelen onderzocht worden; handelsvolume en bid-ask spread. Voor deze metrische eenheden is gekozen vanwege het gemak en de eenvoud van het meten. Allereerst wordt gekeken naar het dagelijkse handelsvolume van de Tesla-aandelen. Er zal een enkelvoudige regressieanalyse worden gemaakt. De regressieanalyse zal de volgende vorm hebben:

$$\text{Volume} = \alpha + \beta T + u$$

Vervolgens wordt de correlatie met de bid-ask spread en de hoeveelheid tweets onderzocht met een regressieanalyse. Deze heeft de volgende vorm:

$$\text{Bid} - \text{Ask Spread} = \alpha + \beta T + u$$

## Effect van Elon Musk's Twitter-gebruik

Hierbij is de bid-ask spread berekend door middel van de bid sluitprijs en ask sluitprijs. Vervolgens wordt de bid-ask spread gedeeld door de aandelenkoers, zodat er over de tijd een relatieve relatie wordt waargenomen tussen bid-ask spread en de aandelenprijs. Wanneer er een significant effect gevonden wordt tussen de afhankelijke variabele en de onafhankelijke variabele (T) wordt, onder de nulhypothese van geen effect,  $H_0$  verworpen.

$$H_0: \beta = 0$$

$$H_a: \beta \neq 0$$

In deze regressies wordt de foutenterm ook robuust gemaakt en worden dezelfde controlevariabelen gebruikt als bij hypothese 1 (zie 3.1).

### 3.3 Hypothese 3

In het geval van correlatie, zoals onderzocht in hypothesen 1 en 2, zal er gekeken worden of er ook sprake is van een 'oorzaak en gevolg' door de hoeveelheid tweets. Correlatie staat namelijk niet gelijk aan causaliteit (Gardner, 2000). De causaliteit zal worden onderzocht door middel van een 'Granger causality' test, met deze test wordt bekeken of de vertragingen (*lags*) voorspellende kracht bevatten over de afhankelijke variabele. Hiervoor wordt eerst een vector autoregressie (VAR) model met een aantal vertragingen gemaakt. Dit model wordt gebruikt om tijdreeks variabelen te schatten. De vertragingen zijn in dit model de regressoren. Het aantal vertragingen wordt bepaald aan de hand van de stata command `varsoc`.

$$H_0: X \text{ does not Granger cause } Y$$

$$H_a: X \text{ Granger causes } Y$$

## Hoofdstuk: 4 Data

### 4.1 Begrippen

In deze paragraaf worden de begrippen die in de hypothese zijn gebruikt (volatiliteit, handelsvolume, bid-ask spread en liquiditeit) toegelicht. Het eerste begrip, volatiliteit, kent geen concrete definitie. Intuïtief kan volatiliteit uitgelegd worden als de mate waarin een aandeel omhoog en omlaag beweegt in een korte periode. Hoe meer een aandeel fluctueert, hoe groter de volatiliteit. Dit betekent dat naarmate de volatiliteit toeneemt, de onzekerheid over een aandeel groeit. In zekere zin kan volatiliteit gezien worden als risico van een investering. Met handelsvolume worden alle transacties die op één dag zijn voltooid, bedoeld. De bid-ask spread bestaat uit de som van de koop premium en de verkoop concessie (Amihud & Mendelson, Asset pricing and the bid-ask spread, 1986). Een grotere bid-ask spread heeft een positief effect op de koerswinst die behaald kan worden. Dit heeft te maken met een groter risico vanwege de handelskosten (Amihud & Mendelson, Liquidity and Stock returns, 1986). Dit is in lijn met het positieve verband tussen de volatiliteit van een aandeel en de bid-ask spread (Wang & Yau, 2000). Een nauwere bid-ask spread is een indicatie van een liquide aandeel. Liquide aandelen zijn aandelen waarbij het relatief makkelijk is om een aandeel te kopen of verkopen. Dit wordt veroorzaakt door de lagere kosten en dus een verminderd risico. Een liquide aandeel wordt verkocht met een liquide premium en een illiquide aandeel wordt verkocht met een liquide korting (Chen, Ibbotson, & Hu, 2013). De liquiditeit van een aandeel kan op verschillende manieren worden gemeten, waaronder het handelsvolume, bid-ask spread en de omloopsnelheid van aandelen. Het handelsvolume heeft een positief effect op de liquiditeit; hoe meer aandelen verhandeld zijn op één dag, hoe hoger de liquiditeit (Aitken & Comerton-Forde, 2003). De bid-ask spread heeft een negatief effect op de liquiditeit; hoe groter de bid-ask spread, hoe minder de liquiditeit. Dit wordt veroorzaakt door de hogere kosten bij een grotere bid-ask spread. De omloopsnelheid heeft een positief effect op de liquiditeit; hoe sneller aandelen van investeerder wisselen hoe hoger de liquiditeit.

### 4.2 Twitteraccount Elon Musk

Het Twitter-account van Elon Musk (@ElonMusk) is opgericht in juni 2009. De eerste tweet vond een jaar later plaats in juni 2010. Voor dit onderzoek zijn Musk's tweets van juni 2010 tot en met april 2020 gebruikt. Alle tweets zijn gegenereerd via Python 3.8 Er is gebruik gemaakt van de python *library* 'tweepy' en een Twitter 'developer' account. (zie appendix III – code 1) door middel van een Twitter API, een publiek dataplatform. Voor de pythoncode is gebruik gemaakt van verschillende bronnen (Soule, 2017), (Rastogi, sd), (Tweepy, 2020). De hoeveelheid tweets is vastgesteld met behulp van Excel.

#### 4.2.1 Twitter-gebruik Elon Musk

Elon Musk bezit twee Twitter-accounts. Het algemene account @ElonMusk en @BoredElonMusk. Het tweede account wordt gebruikt voor niet-relevante gedachtes en niet-bestaande uitvindingen. Zelf noemt Musk @BoredElonMusk het zijn parodie-account. Het account bevat geen informatie over Tesla Inc., maar ideeën in de trant van Musk, Elon (BoredElonMusk). "The [COVID-19] vaccine should be named Norona" 10 Maart 2020, 18:30. Tweet. Dit account zal daarom niet meegenomen worden in het onderzoek.

Elon Musk is tien jaar actief op zijn Twitter-account @ElonMusk. Op dit account heeft hij meer dan 5000 tweets en bijna 40 miljoen volgers. Dit maakt hem één van de meest actieve Twitter-gebruikers onder de CEO's van de *Fortune 500* (Malholtra & Malholtra, 2016). De tweets van Elon Musk bestaan uit drie soorten: tweets, mentions, en ReTweets. Een 'mention' is het benoemen van iemand anders zijn account in een Twitter bericht (door middel van iemands @). Hiermee kan een berichtje naar iemand gestuurd worden of een tweet van iemand beantwoord worden. Een 'ReTweet' is het plaatsen of overnemen van iemand anders zijn Twitter bericht op een eigen account. Volgens een artikel van Cha et al. (2010) bevatten ReTweets en mentions van populaire gebruikers weinig relevante informatie. Musk deelt echter wel vaak nieuwe informatie in zijn mentions, daarnaast nemen de mentions meer dan de helft (59,13%) van zijn gehele Twitter-gebruik in beslag (Tweetstat, 2020). Daarom is besloten om deze soort Twitter berichten wel mee te nemen. De ReTweets zijn buiten beschouwing gelaten. In tabel 1 is een deel van de gebruikte data verkregen uit het Twitteraccount @ElonMusk te zien.

Tabel 1: Voorbeeld dataset tweets

| id                 | Datum          | Tweet  |
|--------------------|----------------|--|
| 849636868052275000 | 5-4-2017 14:56 | 'And so the robots spared humanity ...'                                      |
| 848943072423497000 | 3-4-2017 16:59 | '@waltmossberg @mims @defcon_5 Et tu, Walt?'                                 |
| 848935705057280000 | 3-4-2017 16:30 | 'Stormy weather in Shortville ...'   |
| 848416049573658000 | 2-4-2017 06:05 | "@DaveLeeBBC @verge Coal is dying to nat gas fracking. It's basically dead." |

*Noot.* Het id is een unieke code per tweet

Een tekortkoming van deze methode is dat Twitter API alleen de afgelopen 3200 tweets kan genereren. Dit leidt tot een dataset met tweets van 13-06-2019 tot 14-04-2020. Via data.world is een dataset gevonden die op dezelfde manier gegeneerd is (data.World, 2017). Deze set bevat de tweets van 4-6-2010 tot 5-4-2017. De data tussen april 2017 en juni 2019 is met de hand geteld. In totaal zijn er 6020 tweets geobserveerd.



## Effect van Elon Musk's Twitter-gebruik

| Datum     | Tweet  | Compound | Negatief | Neutraal | Positief |
|-----------|--|----------|----------|----------|----------|
| 18-5-2020 | '@SciGuySpace Zach made a significant contribution to SpaceX'                            | 0.2023   | 0        | 0.816    | 0.184    |
| 17-5-2020 | '@flcnhvy @bluemoondance74 @jeff_foust Great. Just sleeps & eats, overall pretty chill.' | 0.4215   | 0        | 0.682    | 0.318    |

*Noot.* De compound is de polariteit waarde. Het is berekend door de som te nemen van alle waardes die gegeven zijn per woord in de tweet. Vervolgens is de som aangepast aan een aantal regels en genormaliseerd naar een waarde tussen -1 en 1.

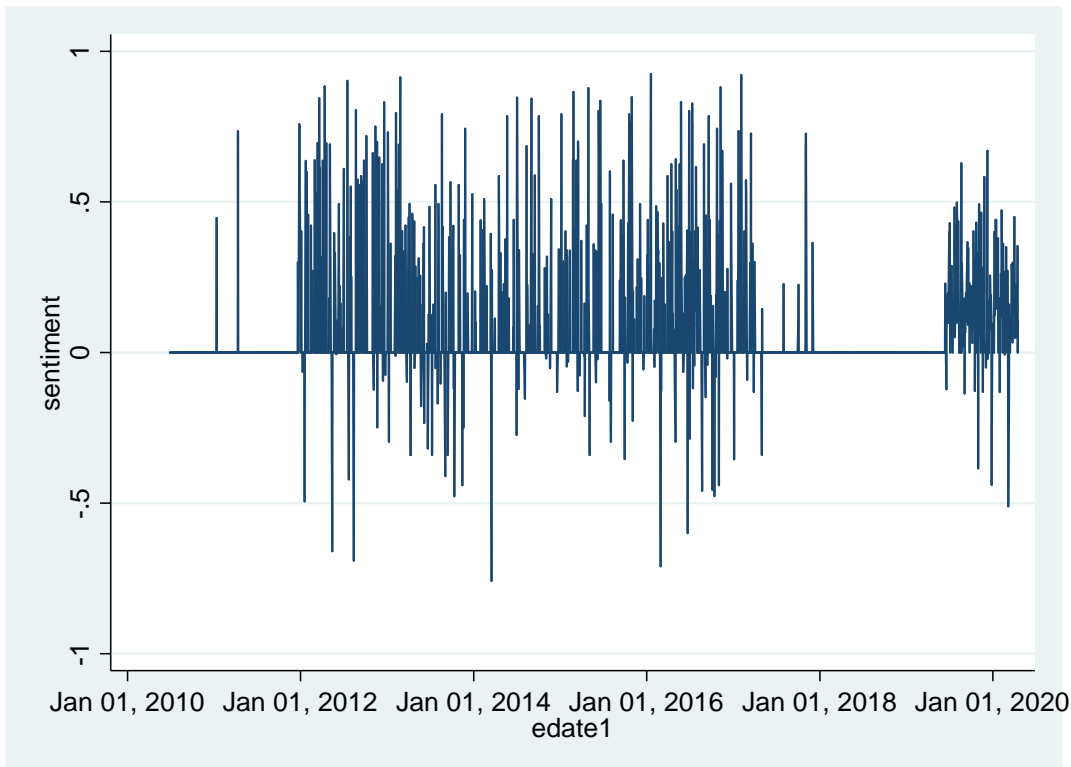
In tabel 2 zijn een aantal tweets met het bijbehorende sentiment uit mei 2020 te zien. De waardes in de kolommen geven aan hoe negatief of positief de woorden in de tweets zijn. Hierbij betekent een nul in de kolom negatief dat de tweet geen negatieve woorden bevat. De drie kolommen samen tellen altijd op tot een waarde van één. Wanneer zowel de kolom negatief als de kolom positief een nul bevat, bevat de kolom neutraal dus automatisch een één. De tweet in tabel 3 bevat een negatieve waarde van 0,351. In dit geval wordt 'too' (high) opgevat als negatief. De overige woorden bevatten geen sentiment ( $0,351 + 0,649 = 1$ ). De tweet is dus 35,1% negatief en 64,9% neutraal. Uit tabel 2 en tabel 3 is vooral de kolom met compound belangrijk. Dit geeft de samengevatte sentiment waarde van de tweet weer. Hierbij wordt een score tussen -0,05 en 0,05 aangemerkt als 'neutraal'. Groter dan 0,05 is positief en kleiner dan -0,05 is negatief. De tweet in tabel 3 wordt dus gezien als een 'negatieve' tweet ( $-0,4019 < -0,05$ ).

Tabel 3: Tweet uitgelicht uit VADER output

| Datum      | Tweet                               | Compound | Negatief | Neutraal | Positief |
|------------|-------------------------------------|----------|----------|----------|----------|
| 30-04-2020 | 'Tesla stock price is too high imo' | -0,4019  | 0,351    | 0,649    | 0        |

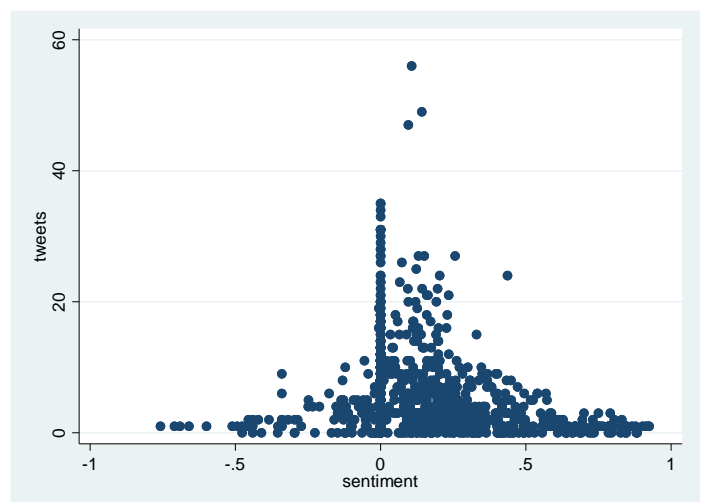


## Effect van Elon Musk's Twitter-gebruik



Afbeelding 2: Twitter-sentiment @ElonMusk, 2010-2020

In afbeelding 2 is te zien dat het sentiment vaker aan de positieve kant zit, dan aan de negatieve kant. Ook is te zien dat het sentiment vanaf 2018 geconcentreerder is tussen de 0 en 0,5; er zijn minder hoge pieken en dalen dan voor 2018. Dit kan betekenen dat de tweets minder gevoel bevatten, dus meer neutraal zijn. In het puntdiagram in afbeelding 3 is te zien dat het sentiment, afgezet tegen de hoeveelheid tweets, links scheef is. Er zijn meer positieve observaties dan negatieve observaties. Het is ook zichtbaar dat meer tweets op één dag leiden tot een neutraler gemiddeld sentiment op diezelfde dag. Wanneer naar gemiddeld sentiment wordt gekeken, is de periode na 2018 inderdaad neutraler. Het gemiddelde sentiment na 2018 is 0,15 tegen een gemiddelde van 0,20 voor 2018. Dit kan een gevolg zijn van de schikking die Elon Musk heeft getroffen met de SEC in 2018. Het gemiddelde sentiment over de gehele periode is 0.174051. Dit betekent dat de gemiddelde tweet een positief gevoel bevat. De meest positieve tweet



Afbeelding 3: puntdiagram sentiment

## Effect van Elon Musk's Twitter-gebruik

heeft een waarde van 0,9558 en de meest negatieve tweet heeft een waarde van -0,8779 (zie tabel 3).

Tabel 3: Meest positieve tweet en meest negatieve tweet

| Datum      | Tweet   | Compound | Negatief | Neutraal | Positief |
|------------|---|----------|----------|----------|----------|
| 10-9-2015  | Yo, I don't hate Apple. It's a great company with a lot of talented people. I love their products and I'm glad they're doing an EV.   | 0.9558   | 0        | 0.544    | 0.456    |
| 17-03-2020 | @SciGuySpace Yeah. There was also an early engine shutdown on ascent, but it didn't affect orbit insertion. Shows value of having 9 engines! Thorough investigation needed before next mission. | -0.8779  | 0.373    | 0.627    | 0        |

### 4.3 Variabelen

Data over de aandelen van Tesla is afkomstig van Yahoo Finance. Hier kan historische data gedownload worden, waaronder openkoers, aangepaste sluitkoers, hoogste en laagste koers per dag en handelsvolume. Met behulp van de openkoers en aangepaste sluitkoers wordt de *return* berekend.

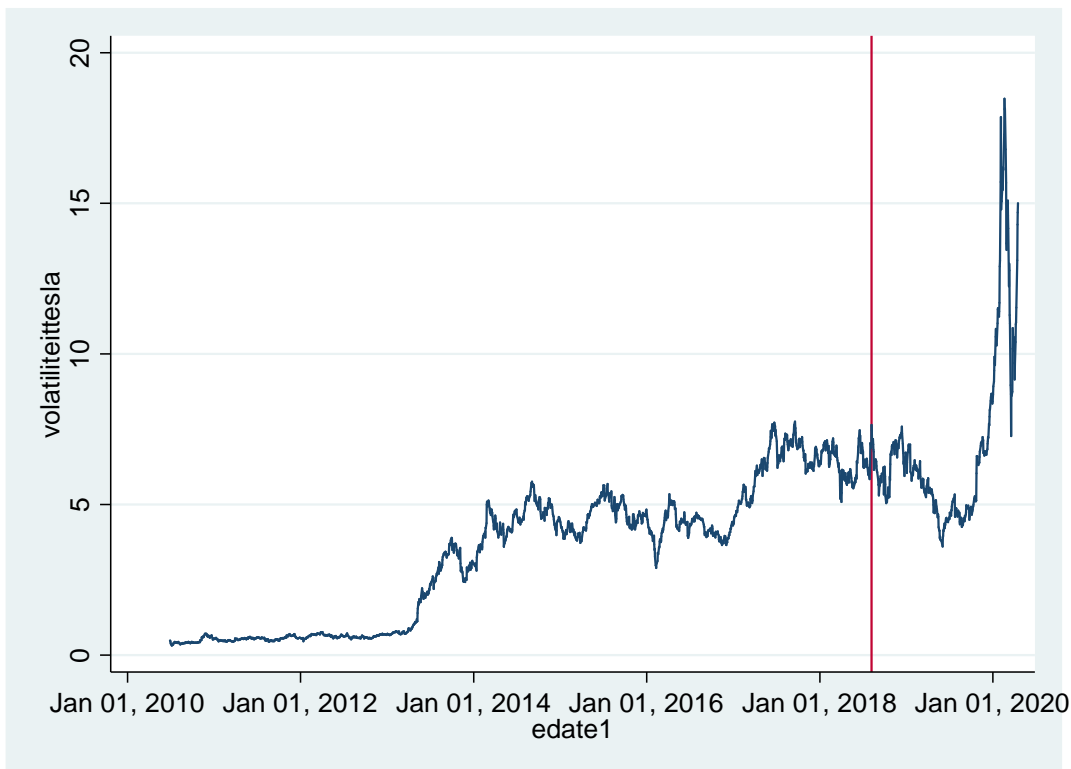
$$\text{Return: } (Aangepaste\ sluitkoers / openkoers) - 1$$

Vervolgens wordt, zoals eerder genoemd in 4.1, aan de hand van de return de variantie ( $\sigma^2$ ) en volatiliteit berekend.

$$\text{Variantie } (\sigma^2) : (X - \mu)^2 / N$$

$$\text{Volatiliteit } (\sigma) : \sqrt{\sigma^2}$$

De volatiliteit is aangegeven in afbeelding 4. In deze grafiek is ook een verticale lijn weergegeven op de datum waarop Elon Musk de \$420-tweet heeft geplaatst.



Afbeelding 4: Volatiliteit Tesla, 2010-2020

De bid-ask spread is berekend aan de hand van Tesla's dagelijkse sluit bid en sluit ask. Vervolgens is de bid-ask spread gedeeld door de aandelenkoers, zodat er over de tijd een relatieve relatie wordt waargenomen tussen bid-ask spread en de aandelenprijs. De bid-ask spread data is verkregen via WRDS uit de CRSP *Daily Stock* (2020). Andere belangrijke controlevariabelen zijn de *debt to equity* ratio van Tesla en de marktvolatiliteit. De D/E ratio is afkomstig van MacroTrends (Tesla Debt to Equity Ratio

2009-2020 | TSLA, 2020). De marktvolatiliteit is berekend aan de hand van dagelijkse NASDAQ-returns, data afkomstig van Yahoo Finance. De hoeveelheid tweets is afkomstig van Elon Musk's persoonlijke Twitter account @ElonMusk.

### 4.3.1 Stationariteit

Om de hypothesen te beantwoorden, zal er gebruik gemaakt worden van een lineaire regressie. Hierbij wordt tijdreeks data gebruikt. Om een betrouwbare schatting te kunnen maken, moeten er aan een aantal voorwaarden voldaan zijn. Hierbij wordt eerst gekeken of de data stationair is en na het uitvoeren van de regressie zal gecontroleerd worden op autocorrelatie. Om te controleren of de variabelen onafhankelijk bewegen van tijd, ofwel stationair zijn, wordt er een Dickey-Fuller test uitgevoerd op de variabelen.

$H_0$ : Unit Root aanwezig

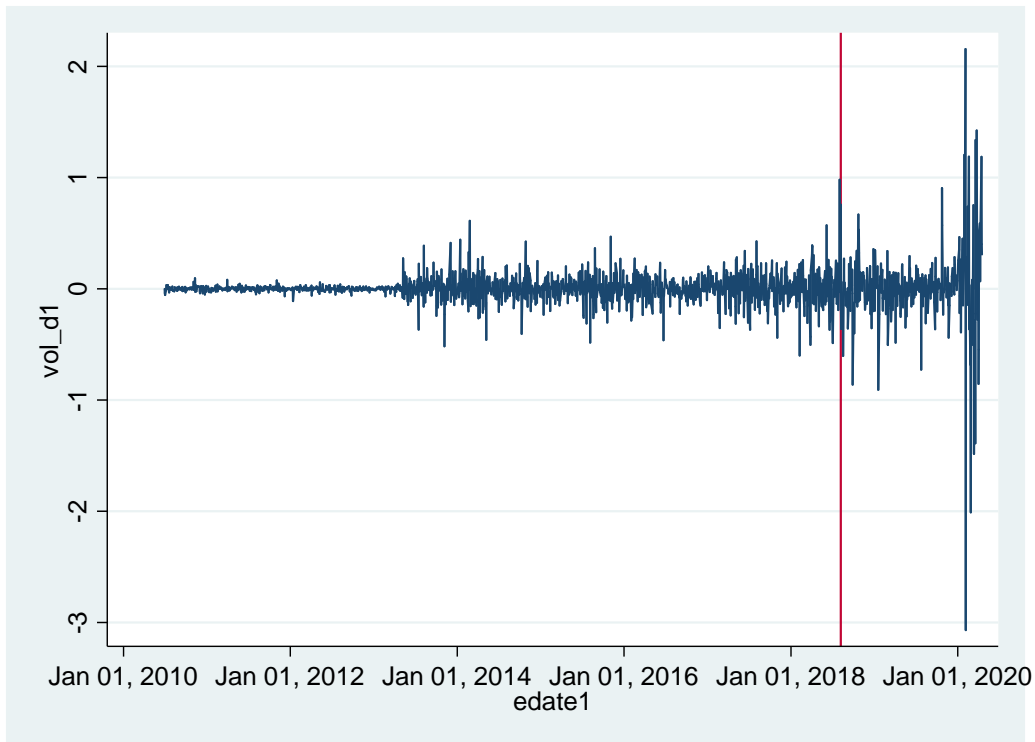
$H_a$ : Unit Root afwezig, dus stationariteit

De test voor de variabele volatiliteit laat zien dat de nulhypothese niet verworpen kan worden, dus er is sprake van niet stationair (zie Appendix II, tabel G). Deze uitkomst is ook te zien in de lijngrafiek, de volatiliteit stijgt over tijd, dus er is sprake van een positieve trend (zie afbeelding 4). De variabele volatiliteit wordt stationair gemaakt door middel van de eerste verschillen:

$$(W_t = Y_t - Y_{t-1}).$$

$W_t$  staat voor de stationaire variabele. Door de eerste verschillen te nemen, wordt de positieve trend uit de variabele gehaald. Na deze transformatie van de variabele wordt de Dickey-Fuller test opnieuw uitgevoerd, hieruit blijkt dat de variabele nu wel stationair is (zie afbeelding 5).

## Effect van Elon Musk's Twitter-gebruik



Afbeelding 5: Volatiliteit stationair

De test voor de variabele D/E laat zien dat de nulhypothese niet verworpen kan worden, dus ook deze variabele is niet stationair (zie appendix II, tabel H). Deze variabele wordt op dezelfde manier stationair gemaakt als de variabele volatiliteit. Bij de overige variabelen kan de nulhypothese verworpen worden, dit betekent dat alle andere variabelen stationair zijn.

## Hoofdstuk 5: Resultaten

### 5.1 Hypothese 1

Hypothese 1 luidt als volgt:

*De hoeveelheid Tweets van Elon Musk is positief gecorreleerd met de volatiliteit van Tesla-aandelen.*

Aan de hand van een lineaire regressieanalyse zal deze correlatie worden getest. Deze regressie ziet er als volgt uit:

$$\sigma_{e_i} = \alpha + \beta T + u$$

De nulhypothese en alternatieve hypothese luiden als volgt:

$$H_0: \beta = 0$$

$$H_a: \beta \neq 0$$

Om te controleren of er geen sprake is van autocorrelatie met de fouteterm in dit model, wordt er een Durbin-Watson test uitgevoerd:

$$H_0: \text{Residuals niet gecorreleerd}$$

$$H_a: \text{Autocorrelatie}$$

$$\text{Durbin - Watson } d - \text{ statistic}(4, 1929) = 1.685693$$

De kritieke waarde in dit model is [1,892;1,90]. De d-waarde van de test buiten deze waardes. Dit betekent dat  $H_0$  wordt verworpen, er wordt positieve autocorrelatie waargenomen. Autocorrelatie betekent dat de data onderling met elkaar gecorreleerd is, dit leidt tot een afwijking in de OLS-schatter. Dit wordt opgelost aan de hand van een regressie met Newey-West standaardfout. Newey en West (1987) hebben voor tijdreeks data een OLS-schatter ontworpen die ook consistent is in de aanwezigheid van zowel heteroskedasticiteit als autocorrelatie. De coëfficiënten worden dus OLS geschat, maar er worden 'Newey West standard errors' gebruikt. Deze methode vraagt om een maximale lengte van de vertragingen om autocorrelatie vast te stellen. De gebruikte formule is als volgt (Brooks, 2019):

$$\text{INTEGER}[4 * (T/100)^{(2/9)}]$$

De regressie met Newey-West standaardfout wordt uitgevoerd met acht vertragingen. Uit de test volgt dat de variabele 'tweets' een positieve coëfficiënt heeft, maar geen significant verklarend effect heeft op de volatiliteit ( $p = 0.324$ ) (zie tabel 5). De controlevariabelen die ook insignificant blijken

## Effect van Elon Musk's Twitter-gebruik

(D/E, bid-ask spread en handelsvolume) worden uit het model verwijderd variabelen. De variabele 'tweets' blijkt nog steeds niet significant. De nulhypothese wordt verworpen.

Tabel 5: Lineaire regressiemodel met Newey-West standaardfout

| Volatiliteit | Coëfficiënt | Newey West Std. Err. | t     | P> t  | [95% Conf. Interval] |          |
|--------------|-------------|----------------------|-------|-------|----------------------|----------|
| Tweets       | .00090      | .00091               | 0.99  | 0.324 | -.000889             | .0026854 |
| _cons        | -.0006      | .00279               | 0.820 | 0.820 | -.0060986            | .004829  |

Hypothese 1 luidde: *'De hoeveelheid Twitter berichten van Elon Musk is positief gecorreleerd met de volatiliteit van Tesla-aandelen'*. Hoewel er een positief effect wordt waargenomen, kan deze hypothese niet worden aangenomen. Er is geen significant bewijs gevonden, dat de hoeveelheid Twitter berichten gecorreleerd is aan de volatiliteit.

## 5.2 Hypothese 2

Hypothese 2 luidt als volgt:

*De hoeveelheid Tweets van Elon Musk is positief gecorreleerd met de liquiditeit van Tesla-aandelen.*

De hypothese wordt getoetst aan de hand van twee variabelen: het handelsvolume en de bid-ask spread van het Tesla aandeel.

### 5.2.1 Handelsvolume

Aan de hand van een lineaire regressieanalyse zal deze correlatie worden getest. Deze regressie ziet er als volgt uit:

$$Volume = \alpha + \beta T + u$$

De nulhypothese en alternatieve hypothese luiden als volgt:

$$H_0: \beta = 0$$

$$H_a: \beta \neq 0$$

Om te controleren of er geen sprake is van autocorrelatie met de foutenterm in dit model, wordt er een Durbin-Watson test uitgevoerd.

$$H_0: \text{Residuals niet gecorreleerd}$$

$$H_a: \text{Autocorrelatie}$$

$$\text{Durbin - Watson } d - \text{ statistic}(2, 2467) = .4331476$$

De kritieke waarde in dit model is [1,63;1,72]. De d-waarde van de test valt buiten deze waardes. Dit betekent dat  $H_0$  verworpen kan worden; er wordt autocorrelatie waargenomen. Dit betekent dat de data onderling met elkaar gecorreleerd is, wat leidt tot een afwijking in de OLS-schatter. De regressie met Newey-West standaardfout wordt uitgevoerd met acht vertragingen. Uit de test volgt dat de variabele 'tweets' een significant positief verklarend effect heeft op het handelsvolume ( $p = 0.0$ ) (zie tabel 6). De nulhypothese van geen effect wordt verworpen, er wordt geconcludeerd dat de hoeveelheid tweets een positief effect heeft op het handelsvolume van Tesla. Uit de tabel kan verder nog worden opgemaakt dat marktvolatiliteit een significant positief effect heeft op het handelsvolume, de overige controlevariabelen bleken niet significant.

Tabel 6: Lineaire regressiemodel met Newey-West standaardfout

| Handels-<br>volume    | Coëfficiënt | Newey<br>West<br>Std. Err. | t     | P> t  | [95% Conf. Interval] |         |
|-----------------------|-------------|----------------------------|-------|-------|----------------------|---------|
| Tweets                | 352000      | .00091                     | 7.39  | 0.0   | 259000               | 446000  |
| D/E                   | -46000      | .00279                     | -0.77 | 0.441 | -163000              | 71100   |
| Markt<br>volatiliteit | 3480000     | 1240000                    | 2.80  | 0.005 | 1040000              | 5920000 |
| _cons                 | 4342000     | 282000                     | 15.38 | 0.0   | 3790000              | 4900000 |

*Noot.* Afgerond op duizenden

### 5.2.2 Bid-ask spread

Aan de hand van een lineaire regressieanalyse zal de correlatie worden getest. Deze regressie ziet er als volgt uit:

$$\text{Bid-ask spread} = \alpha + \beta T + u$$

Om te controleren of er geen sprake is van autocorrelatie met de foutenterm in dit model, wordt er een Durbin-Watson test uitgevoerd.

$H_0$ : Residuals niet gecorreleerd

$H_a$ : Autocorrelatie

$$\text{Durbin-Watson } d\text{-statistic}(4, 1872) = 1.712355$$

De kritieke waarde in dit model is [1,894;1,898]. De d-waarde van de test valt buiten deze waardes. Dit betekent dat  $H_0$  verworpen kan worden; er wordt autocorrelatie waargenomen. Dit betekent dat de data onderling met elkaar gecorreleerd is, dit leidt tot een afwijking in de OLS-schatter. De regressie met Newey-West standaardfout wordt uitgevoerd met acht vertragingen. Uit de test volgt dat de variabele 'tweets' een negatieve coëfficiënt heeft, maar geen significant verklarend effect heeft op de bid-ask spread ( $p = 0,309$ ) (zie tabel 7). Hoewel er een negatief effect wordt waargenomen op



bid-ask spread, wat leidt tot een grotere liquiditeit, kan deze hypothese niet worden aangenomen. Er is geen significant bewijs gevonden, dat de hoeveelheid Twitter berichten gecorreleerd is aan de bid-ask spread. Uit de tabel kan verder nog worden opgemaakt dat marktvolatiliteit een significant positief effect heeft op de bid-ask spread, dit effect bleek ook uit eerdergenoemd literatuur (Wang & Yau, 2000). De controlevariabele schuldratio blijkt een significant positief effect te hebben en return een significant negatief effect op de bid-ask spread. Het handelsvolume bleek, tegen verwachting in, geen effect te hebben op de bid-ask spread.

Tabel 7: Lineaire regressiemodel met Newey-West standaardfout

| Bid-Ask spread percentage        | Coëfficiënt | Newey West Std. Err. | t     | P> t  | [95% Conf. Interval] |          |
|----------------------------------|-------------|----------------------|-------|-------|----------------------|----------|
| Tweets                           | -.0004      | .0004                | -1.02 | 0.309 | -.0012               | .0004    |
| Handelsvolume Markt volatiliteit | 2.52e-10    | 5.69e-10             | 0.44  | 0.658 | -8.64e-10            | 1.37e-19 |
| D/E                              | .0912       | 0.00403              | 22.63 | 0.0   | .0833                | .0991    |
| Return                           | 0.0016      | 0.0007               | 2.29  | 0.022 | .0002                | .0030    |
| _cons                            | -.3007      | .0001                | -2.68 | 0.007 | -.5207               | -.0807   |
|                                  | 1.00        | .0385                | 26.01 | 0.0   | .9272                | 1.0784   |

Noot. Alle getallen in duizendtallen

### 5.2.3 Liquiditeit

Hypothese 2 luidde: *'De hoeveelheid Twitter berichten van Elon Musk is positief gecorreleerd met de liquiditeit van Tesla-aandelen.'* De liquiditeit is gemeten aan de hand van twee variabelen:

handelsvolume (besproken in 5.2.1) en bid-ask spread (besproken in 5.2.2). Er is gebleken dat het aantal tweets een significant positief effect heeft op het handelsvolume, dit betekent dat een grotere hoeveelheid tweets leidt tot een hoger handelsvolume. Er is geen significant bewijs gevonden dat de hoeveelheid Tweets gecorreleerd is aan de bid-ask spread. De conclusie voor de hypothese is hierdoor onzeker. Het aantal Tweets beïnvloedt handelsvolume. Dit leidt tot een verhoogde liquiditeit, dit is in lijn met eerder benoemde literatuur (Blankespoor, Miller, & Hal, 2014). Het effect is niet waargenomen tussen het aantal Tweets en de bid-ask spread. Mogelijk zijn er een aantal tekortkomingen, waardoor de kracht van de test afneemt. Dit kan veroorzaakt zijn door een weggelaten controlevariabele die van belang is of te weinig observaties, hiervoor is meer onderzoek nodig (zie Aanbevelingen). Tussen de hoeveelheid Tweets en de liquiditeit van Tesla-aandelen wordt een onzeker positieve correlatie waargenomen.

### 5.3 Hypothese 3

Hypothese 3 luidt als volgt:

*De hoeveelheid Tweets van Elon Musk bevat voorspellende waarde over de volatiliteit en liquiditeit.*

De hypothese wordt onderzocht aan de hand van de Granger-causality test. Deze test kent de volgende hypothesen:

$H_0$ : X does not Granger cause Y

$H_a$ : X Granger causes Y

In de Granger causality test is gebruik gemaakt van de variabelen handelsvolume, bid-ask spread en tweets. Hiervoor is gekozen, omdat in de eerder behandelde hypothese significante correlatie is waargenomen tussen handelsvolume – tweets en bid-ask spread - tweets. Uit de STATA commando *varsoc* is gebleken dat drie vertragingen (*lags*) de optimale keus is (zie appendix II – tabel F). Er wordt gekeken of er een vertraagd effect is van tweets die tot drie dagen eerder zijn geplaatst. Uit de test wordt de volgende conclusie getrokken (zie tabel 8).

Tweets Granger cause volume ( $0,001 < 0,05$ ), de nulhypothese wordt verworpen.

Tweets do not Granger cause bid-ask spread ( $0,907 \neq < 0,05$ ), de nulhypothese wordt niet verworpen.

Tabel 8: Belangrijkste resultaten Granger-Causality test

| Equation                  | Excluded | Chi <sup>2</sup> | df | Prob > chi <sup>2</sup> |
|---------------------------|----------|------------------|----|-------------------------|
| Handelsvolume             | Tweets   | 13.6             | 3  | .003                    |
| Bid-ask spread percentage | Tweets   | .55              | 3  | .907                    |

Noot. Volledige resultaten, zie appendix II – tabel G

Er kan geconcludeerd worden dat de hoeveelheid tweets gebruikt kan worden bij het voorspellen van handelsvolume. Dus zoals onderzocht in hypothese 2, correleren tweets niet alleen met het handelsvolume, maar bevatten ze ook een voorspellende waarde. De hoeveelheid tweets kan niet gebruikt worden als voorspellende waarde van de bid-ask spread. Er wordt geconcludeerd dat de hoeveelheid tweets een voorspellende waarden is voor liquiditeit door middel van het handelsvolume.

## Hoofdstuk 6: Conclusie, Discussie en Aanbevelingen

### 6.1 Conclusie

In dit onderzoek is één centrale vraag aan de hand van een literatuuronderzoek en verschillende hypothesen onderzocht. De vraag luidde:

*Leidt het Twitter-gedrag van Elon Musk tot een positief of negatief effect op de prestatie van Tesla Inc. aandelen?*

Uit de verschillende besproken onderzoeken blijkt dat Tweets invloed hebben op investeerders. Niet alleen worden Twitter-accounts gemonitord als nieuwsplatform, maar ook wordt de aandelenmarkt voorspeld aan de hand van Twitter-sentiment. In het geval van Elon Musk wordt het Twitter-account @ElonMusk gebruikt als nieuwsplatform waarop regelmatig aankondigingen worden gedaan over Tesla. Wanneer het Twitter-sentiment van dit account wordt bekeken, wordt er een positief sentiment over de jaren waargenomen. Investeerders worden op Twitter dus geconfronteerd met gemiddeld positieve tweets, wat kan leiden tot een positief effect op de aandelenmarkt (Bollen, Mao, & Zheng, 2011).

Uit de regressieanalyse met Newey-West standaardfout van de eerste hypothese, *De hoeveelheid Tweets van Elon Musk is positief gecorreleerd met de volatiliteit van Tesla-aandelen*, wordt geen significant positief effect waargenomen. Dit betekent dat er niet uitgesloten kan worden dat de hoeveelheid Tweets geen effect heeft op de volatiliteit van Tesla-aandelen. Uit de regressieanalyse van de tweede hypothese, *De hoeveelheid Tweets van Elon Musk is positief gecorreleerd met de liquiditeit van Tesla-aandelen*, wordt een twijfelend beeld geschetst. De liquiditeit wordt gemeten aan de hand van handelsvolume en bid-ask spread. Er is een significant positief effect waargenomen tussen de hoeveelheid tweets en het handelsvolume van Tesla-aandelen, maar tussen de hoeveelheid tweets en de bid-ask spread is geen significant negatief effect waargenomen. De hypothese kan niet met zekerheid worden bevestigd. Wel kan geconcludeerd worden dat meer tweets leiden tot een hoger handelsvolume in de aandelen van Tesla. Volgens eerdergenoemd onderzoek (Blankespoor, Miller, & Hal, 2014) leidt dit tot een verhoging van de liquiditeit van een aandeel. Een negatief effect van tweets op de bid-ask spread, leidt ook tot een verhoging van de liquiditeit, dit is in lijn met (Blankespoor, Miller, & Hal, 2014), maar voor dit effect is geen significant bewijs gevonden. In de derde hypothese wordt deze correlatie tussen tweets en liquiditeit verder onderzocht door middel van een Granger causality test. Hieruit wordt waargenomen dat tweets als voorspellende waarde gebruikt kan worden voor het handelsvolume. Tussen tweets en bid-ask spread werd ditzelfde verband niet waargenomen.

Uit dit onderzoek wordt geconcludeerd dat het gemiddeld positieve sentiment van de tweets kan leiden tot een positief effect op de aandelenkoers van Tesla. Ook is er een positief effect tussen de hoeveelheid tweets en het handelsvolume van Tesla gevonden, maar tussen de hoeveelheid tweets en bid-ask spread is geen significant negatief effect waargenomen. Het is onzeker of tweets een effect hebben op de liquiditeit van Tesla-aandelen, hiervoor is meer onderzoek nodig.

### 6.2 Discussie en aanbevelingen

Een tekortkoming in het onderzoek is het ontbreken van Twitter-data in 2018. Hierdoor is er geen volledig beeld geschetst. Het berekende Twitter-sentiment was hierdoor niet volledig bruikbaar; het was niet geschikt voor een regressieanalyse. Hierdoor is het effect van Twitter-sentiment niet volledig onderzocht, wat leidt tot een incomplete sentimentanalyse. Wanneer de data uit 2018 wordt meegenomen, zou dit leiden tot een grotere steekproefomvang. Hierdoor zou een krachtigere regressieanalyse ontstaan met een meer nauwkeurige schatting van de parameters. Daarnaast zijn er in de statistische testen mogelijk meerdere relevante variabelen buiten beschouwing gelaten. Het gevolg hiervan is een mogelijke afwijking in de schatting van de parameters. Voor verder onderzoek is het interessant om het Twitter-sentiment van Elon Musk verder te onderzoeken in relatie met de aandelen van Tesla en kunnen de modellen uitgebreid worden voor nauwkeurige coëfficiëntschatting. Dit onderzoek kan ook toegepast worden op het Twitter-gebruik van meerdere CEO's. Dit zou kunnen bijdragen aan het maximaliseren van Twitter's nut voor CEO's.

## Bibliografie

### Literatuurbronnen

- Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *The Journal of Financial Economics*, 223-249.
- Amihud, Y., & Mendelson, H. (1986). Liquidity and Stock returns. *Financial Analysts Journal*, 43-48.
- Barber, B. M., & Odean, T. (2008). All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *The Review of Financial Studies*, 785-818.
- Bartov, E., Faurel, L., & Mohanram, P. S. (2018). Can Twitter Help Predict Firm-Level Earnings and Stock. *The Accounting Review*, 25-57.
- Blankespoor, E., Miller, G. S., & Hal, D. W. (2014). The Role of Dissemination in Market Liquidity: Evidence from Firms' Use of Twitter™. *The Accounting Review*, 79-112.
- Bollen, J., Mao, H., & Zheng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 1-8.
- Brooks, C. (2019). Dealing with Autocorrelation. In C. Brooks, *Introductory Econometrics For Finance* (p. 203). Reading, UK: Cambridge University Press.
- Chen, Z., Ibbotson, R. G., & Hu, W. (2013). Liquidity as an Investment Style. *Financial Analysts Journal*, 30-44.
- De Maio, C., Fenza, G., Loia, V., & Parente, M. (2016). Time Aware Knowledge Extraction for microblog summarization. *Information Fusion*, 60-74.
- Elliott, W. B., Grant, S. M., & Hodge, D. F. (2018). Negative News and Investor Trust: The Role of Firm and CEO Twitter Use. *Journal of Accounting Research*, 1483-1519.
- Gardner, R. (2000). Correlation, causation, motivation, and second language acquisition. *Canadian Psychology/Psychologie canadienne* (41), 10-24.
- Gilbert, E., & Hutto, C. (2014). Conference: Proceedings of the Eighth International AAI Conference on Weblogs and Social Media, At Ann Arbor, MI. *VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text* (pp. 1-11). Ann Arbor: ResearchGate.
- Kouloumpis, E., Wilson, T., & Moore, J. (2011). Twitter-sentiment Analysis: The Good the Bad and the OMG! *Twitter-sentiment Analysis*, (p. 4). Barcelona.
- Liu, Y., Kliman-Silver, C., & Mislove, A. (2014). Eighth International AAI Conference on Weblogs and Social Media. *The Tweets They Are a-Changin': Evolution of Twitter Users and Behavior*. Ann Arbor, Michigan.
- Malholtra, C. K., & Malholtra, A. (2016). How CEOs can leverage Twitter. *MIT Sloan Management Review*. 57(2), 73.
- Mao, Y., Wang, B., Wei, W., & Liu, B. (2012, August). Correlating S&P 500 Stocks with Twitter Data. *Proceedings of the First ACM International Workshop on Hot Topics on Interdisciplinary Social Networks Research*, 69-72.

## Effect van Elon Musk's Twitter-gebruik

- Mittal, A., & Goel, A. (2012). *Stock Prediction Using Twitter-sentiment Analysis*. Opgehaald van stanford.edu: <http://cs229.stanford.edu/proj2011/GoelMittal-StockMarketPredictionUsingTwitterSentimentAnalysis>
- Simpson, R. K. (2018). Securities Fraud or Elon being Elon? The Forty Million Dollar Tweet. *Kentucky Law Journal (Online)*.
- Rao, T., & Srivastava, S. (2012). Analyzing stock market movements using Twitter-sentiment analysis. Sul, H., Dennis, A. R., & Yuan, L. (2016). Trading on Twitter: Using Social Media Sentiment to Predict Stock Returns. *Decision Sciences*, 454-488.
- Thornton, D. L. (1987). A NOTE ON THE EFFICIENCY OF THE COCHRANE-ORCUTT. *The Journal of Econometrics*, 369-376.
- Zhang, X., Fuehres, H., & Gloor, P. A. (2011). Predicting Stock Market Indicators Through Twitter. *Procedia - Social and Behavioral Sciences*, 55-62.

## Internetbronnen

- (2020, Juni 15). Opgehaald van Tweetstat: <http://www.tweetstats.com/graphs/elonmusk#tstats>
- Aslam, S. (2020, Februari 10). *Twitter by the Numbers: Stats, Demographics & Fun Facts*. Opgehaald van Omnicore: <https://www.omnicoreagency.com/Twitter-statistics/>
- Clinton, E. (2019, Maart 3). *How Many Fortune 500 CEOs are on Social Media?* Opgehaald van Influential Executive: <https://influentialexecutive.com/fortune-500-ceos-on-social/>
- data.World. (2017). *Elon Musk's tweets from 2010-06-04 to 2017-04-05*. Opgehaald van data.world: <https://data.world/adamhelsinger/elon-musk-tweets-until-4-6-17>
- Fundamentals, A. C. (2020, Juli 28). Opgehaald van WRDS: <https://wrds-web.wharton.upenn.edu/wrds/>
- Gardner, R. (2000). Correlation, causation, motivation, and second language acquisition. *Canadian Psychology/Psychologie canadienne* (41), 10-24.
- Kelleher, K. (2018, September 27). *Tesla Stock Plunges After SEC Sues Elon Musk for 'False and Misleading' Statements about Taking Tesla Private*. Opgehaald van Fortune: <https://fortune.com/2018/09/27/sec-elon-musk-lawsuit-tesla-stock/>
- Kumar, N. (2017). *Twitter-sentiment Analysis using Python*. Opgehaald van GeeksforGeeks: <https://www.geeksforgeeks.org/Twitter-sentiment-analysis-using-python/>
- Levy, A. (2018, Augustus 8). *Tesla has its busiest trading day since 2014 after Musk tweets he plans to take company private*. Opgehaald van CNBC: <https://www.cnbc.com/2018/08/07/tesla-has-busiest-trading-day-since-2014.html>
- Mittal, A., & Goel, A. (2012). *Stock Prediction Using Twitter-sentiment Analysis*. Opgehaald van stanford.edu: <http://cs229.stanford.edu/proj2011/GoelMittal-StockMarketPredictionUsingTwitterSentimentAnalysis>
- Money. (2018, augustus 18). *These Nine Words Helped Make Elon Musk \$1.4 Billion Richer in One Day*. Opgehaald van Money: <https://money.com/elon-musk-tweet-tesla-private-net-worth/>

## Effect van Elon Musk's Twitter-gebruik

- Mueller, A. (2020). *WordCloud for Python documentation*. Opgehaald van GitHub: [https://amueller.github.io/word\\_cloud/index.html](https://amueller.github.io/word_cloud/index.html)
- Musk, E. (2018, Augustus 7). *Taking Tesla Private*. Opgehaald van Tesla: [https://www.tesla.com/en\\_GB/blog/taking-tesla-private?redirect=no](https://www.tesla.com/en_GB/blog/taking-tesla-private?redirect=no)
- O'Keane, S. (2019, April 30). *The court has approved Elon Musk's new agreement to let lawyers oversee his Tesla tweets*. Opgehaald van The Verge: <https://www.theverge.com/2019/4/26/18484751/elon-musk-sec-fraud-tesla-tweets-contempt-agreement>
- Pandey, P. (2018, September 23). *Simplifying Sentiment Analysis using VADER in Python (on Social Media Text)*. Opgehaald van Medium: <https://medium.com/analytics-vidhya/simplifying-social-media-sentiment-analysis-using-vader-in-python-f9e6ec6fc52f>
- Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., & Mozetič, I. (2015, September 21). *The Effects of Twitter-sentiment on Stock Price Returns*. Opgehaald van Plos one: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0138441>
- Rastogi, P. (sd). *Extracting Tweets Using Twitter Premium Search API and Python*. Opgehaald van Medium: <https://medium.com/swlh/extracting-tweets-using-Twitter-premium-search-api-and-python-2d025144e8a4>
- SEC. (2013, april 2). *SEC Says Social Media OK for Company Announcements if Investors Are Alerted*. Opgehaald van Sec.gov: <https://www.sec.gov/news/press-release/2013-2013-51htm>
- SEC. (2018, September 27). *Elon Musk Charged With Securities Fraud for Misleading Tweets*. Opgehaald van SEC: <https://www.sec.gov/news/press-release/2018-219>
- SEC. (2018, September 29). *Elon Musk Settles SEC Fraud Charges; Tesla Charged With and Resolves Securities Law Charge*. Opgehaald van SEC: <https://www.sec.gov/news/press-release/2018-226>
- Soule, G. (2017, Februari 7). *Instantly share code, notes, and snippets*. Opgehaald van GitHub: <https://gist.github.com/gabrielsoule/638201ac0cc12828d3cde69035a25336>
- Tesla Debt to Equity Ratio 2009-2020 | TSLA*. (2020, April). Opgehaald van Macrotrends: <https://www.macrotrends.net/stocks/charts/TSLA/tesla/debt-equity-ratio>
- Tweepy*. (2020). Opgehaald van Tweepy.org
- Twitter. (2020, Juli 29). *Hashtags gebruiken*. Opgehaald van Twitter: <https://help.Twitter.com/nl/using-Twitter/how-to-use-hashtags>
- Twitter Developer. (2020). *Twitter IDs (snowflake)*. Opgehaald van Twitter Developer: <https://developer.Twitter.com/en/docs/basics/Twitter-ids>
- wordcloud 1.7.0*. (2020, Mei 2). Opgehaald van Python: <https://pypi.org/project/wordcloud/>

## Appendix

### Appendix I – STATA Do-File

Do file Time Series

*#tijdreeks*

browse

gen edate1=date(Date,"MDY")

format %tdMon\_DD,\_CCYY edate1

tsset edate1, daily

tsfill

*#stationariteit*

dfuller volatilitettesla , regress

line volatilitettesla edate1 , tline(07Aug2018)

dfuller tweets, regress

dfuller volumetesla , regress

dfuller nasdaqreturn, regress

dfuller returnTesla, regress

dfuller DE , regress

line DE edate1

generate vol\_d1 = d1.volatilitettesla

dfuller vol\_d1 , regress

generate de\_d1 = d1.DE

dfuller de\_d1 , regress

*#regressie hypothese 1*

reg vol\_d1 tweets DE\_d1 volumetesla, robust

*#autocorrelatie*

estat dwatson

newey vol\_d1 tweets DE\_d1 volumetesla , lag(8) force

newey vol\_d1 tweets, lag(8) force

*#regressie, hypothese 2*

reg volumetesla tweets vol\_d1 de\_d1 nasdaqreturn returtesla , robust

reg Baperc tweets volumetesla DE\_d1, robust



## Effect van Elon Musk's Twitter-gebruik

### *#autocorrelatie*

estat dwatson

newey volumetesla tweets Baperc DE\_d1 mvolatility\_d1 returtesla nasdaqreturn vol\_d1, lag(8)  
force

newey volumetesla tweets DE\_d1 mvolatility, lag(8) force

newey Baperc tweets volumetesla mvolatility\_In DE\_d1 returtesla, lag(8) force

### *#Granger causaliteit test*

varsoc volumetesla tweets

varsoc Baperc tweets

var volumetesla tweets Baperc, lags(1/3)

varGranger

Appendix II – STATA output

**Tabel A: Hypothese 1, regressie met Newey-West standaardfout**

```
. newey vol_d1 tweets DE_d1 volumetesla , lag(8) force
```

```
Regression with Newey-West standard errors   Number of obs   =   1,872
maximum lag: 8                               F( 3,  1868) =   0.48
                                           Prob > F       =   0.6996
```

|             | Newey-West |           |       |       |                      |          |
|-------------|------------|-----------|-------|-------|----------------------|----------|
| vol_d1      | Coef.      | Std. Err. | t     | P> t  | [95% Conf. Interval] |          |
| tweets      | .0009784   | .0008667  | 1.13  | 0.259 | -.0007213            | .0026781 |
| DE_d1       | .0163354   | .0417152  | 0.39  | 0.695 | -.0654779            | .0981488 |
| volumetesla | -3.12e-10  | 1.77e-09  | -0.18 | 0.860 | -3.79e-09            | 3.17e-09 |
| _cons       | .0008618   | .0075306  | 0.11  | 0.909 | -.0139075            | .0156311 |

**Tabel B: Hypothese 1, regressie met Newey-West standaardfout insignificante controlevariabelen verwijderd**

```
. newey vol_d1 tweets, lag(8) force
```

```
Regression with Newey-West standard errors   Number of obs   =   1,872
maximum lag: 8                               F( 1,  1870) =   0.97
                                           Prob > F       =   0.3244
```

|        | Newey-West |           |       |       |                      |          |
|--------|------------|-----------|-------|-------|----------------------|----------|
| vol_d1 | Coef.      | Std. Err. | t     | P> t  | [95% Conf. Interval] |          |
| tweets | .0008982   | .0009113  | 0.99  | 0.324 | -.000889             | .0026854 |
| _cons  | -.0006348  | .0027859  | -0.23 | 0.820 | -.0060986            | .004829  |

**Tabel C: hypothese 2 regressie met Newey-West standaardfout**

## Effect van Elon Musk's Twitter-gebruik

newey volumetesla tweets Baperc DE\_d1 mvolatility\_d1 returtesla nasdaqreturn vol\_d1, lag(8)  
force

Regression with Newey-West standard errors    Number of obs    =    1,701  
maximum lag: 8                                    F( 7, 1693) =    7.33  
    Prob > F        =    0.0000

---

|                | Newey-West |           |       |       |                      |          |
|----------------|------------|-----------|-------|-------|----------------------|----------|
|                | Coef.      | Std. Err. | t     | P> t  | [95% Conf. Interval] |          |
| tweets         | 262878.7   | 39657.41  | 6.63  | 0.000 | 185096               | 340661.4 |
| Baperc         | -7.54e+08  | 1.49e+09  | -0.51 | 0.614 | -3.68e+09            | 2.18e+09 |
| DE_d1          | -32101.48  | 58914.07  | -0.54 | 0.586 | -147653.5            | 83450.58 |
| mvolatility_d1 | 9.12e+08   | 8.02e+08  | 1.14  | 0.255 | -6.60e+08            | 2.48e+09 |
| returtesla     | 4188182    | 8087293   | 0.52  | 0.605 | -1.17e+07            | 2.01e+07 |
| nasdaqreturn   | 1.86e+07   | 1.10e+07  | 1.69  | 0.092 | -3014393             | 4.01e+07 |
| vol_d1         | -367762.8  | 345880.6  | -1.06 | 0.288 | -1046161             | 310635.8 |
| _cons          | 4896243    | 317504.8  | 15.42 | 0.000 | 4273500              | 5518986  |

---

**Tabel D: hypothese 2 regressie met Newey-West standaardfout insignificante controlevariabelen verwijderd**

. newey volumetesla tweets DE\_d1 mvolatility, lag(8) force

Regression with Newey-West standard errors    Number of obs    =    1,748  
maximum lag: 8                                    F( 3, 1744) =    18.96  
    Prob > F        =    0.0000



Effect van Elon Musk's Twitter-gebruik

Sample: 18459 - 21903, but with gaps      Number of obs    =    409

```

+-----+
|lag | LL  LR  df  p   FPE   AIC   HQIC  SBIC  |
|---+-----|
| 0 | -8122.67                6.2e+14  39.7295  39.7372  39.7491 |
| 1 | -7909.05  427.25  4  0.000  2.2e+14  38.7044  38.7277  38.7633* |
| 2 | -7899.53  19.032  4  0.001  2.2e+14  38.6774  38.7163  38.7756 |
| 3 | -7889.26  20.541  4  0.000  2.1e+14  38.6468  38.7011*  38.7842 |
| 4 | -7884.45  9.6348*  4  0.057  2.1e+14*  38.6428*  38.7127  38.8194 |
+-----+

```

Endogenous: volumetesla tweets

Exogenous: \_cons

. varsoc BAspread tweets

Selection-order criteria

Sample: 18459 - 21903, but with gaps      Number of obs    =    409

```

+-----+
|lag | LL  LR  df  p   FPE   AIC   HQIC  SBIC  |
|---+-----|
| 0 |  -604                .066375  2.96332  2.97109  2.98295 |
| 1 | -541.509  124.98  4  0.000  .049864  2.6773  2.7006  2.73619 |
| 2 | -516.371  50.275  4  0.000  .044968  2.57394  2.61277  2.67208 |
| 3 | -503.839  25.064*  4  0.000  .04313*  2.53222*  2.58658*  2.66961* |
| 4 | -500.863  5.9538  4  0.203  .043347  2.53723  2.60712  2.71387 |
+-----+

```

Endogenous: BAspread tweets

Exogenous: \_cons

**Tabel G – Hypothese 3, Granger causality test**

. vargranger

Granger causality Wald tests

Effect van Elon Musk's Twitter-gebruik

```

+-----+
| Equation      Excluded | chi2  df Prob > chi2 |
+-----+-----+
| volumetesla   tweets   | 18.768  3  0.000 |
| volumetesla   BAspread |  2.7626  3  0.430 |
| volumetesla   ALL      |  24.77   6  0.000 |
+-----+-----+
| tweets        volumetesla | 17.503  3  0.001 |
| tweets        BAspread   |  2.3197  3  0.509 |
| tweets        ALL        | 22.393  6  0.001 |
+-----+-----+
| BAspread      volumetesla |  2.0003  3  0.572 |
| BAspread      tweets     |  4.5088  3  0.212 |
| BAspread      ALL        |  7.5964  6  0.269 |
+-----+-----+

```

**Tabel G: Dickey-Fuller test volatiliteit**

**. dfuller volatiliteitTesla**

Dickey-Fuller test for unit root      Number of obs = 1929

```

----- Interpolated Dickey-Fuller -----
Test   1% Critical  5% Critical  10% Critical
Statistic  Value      Value      Value
-----
Z(t)  -1.039   -3.430   -2.860   -2.570
-----

```

Mackinnon approximate p-value for Z(t) = 0.7388

## Effect van Elon Musk's Twitter-gebruik

**Tabel H: Dickey-Fuller test D/E**

**. dfuller DE**

Dickey-Fuller test for unit root      Number of obs = 1929

----- Interpolated Dickey-Fuller -----

Test    1% Critical    5% Critical    10% Critical

Statistic    Value    Value    Value

-----

Z(t)    -0.607    -3.430    -2.860    -2.570

-----

MacKinnon approximate p-value for Z(t) = 0.8693

### Appendix III – Python codes

#### Code 1: Tweets genereren

```
import tweepy

import csv

#Twitter autoriseren

def get_all_tweets(screen_name):
    consumer_key = "[verwijderd]"
    consumer_secret = "[verwijderd]"
    access_key = "[verwijderd]"
    access_secret = "[verwijderd]"

    auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
    auth.set_access_token(access_key, access_secret)
    api = tweepy.API(auth, wait_on_rate_limit=True)

    #initialiseer 2 lijsten met alle tweets en zonder ReTweets's
    alltweets = []
    noRT = []

    #Verkrijgen tweets
    new_tweets = api.user_timeline(screen_name = screen_name, tweet_mode =
'extended', count=200)

    alltweets.extend(new_tweets)

    oldest = alltweets[-1].id - 1

    while len(alltweets) <= 3200:
        print("getting tweets before {}".format(oldest))

        new_tweets = api.user_timeline(screen_name = screen_name, tweet_mode =
'extended', count=200, max_id=oldest)

        alltweets.extend(new_tweets)

        oldest = alltweets[-1].id - 1

        print("...{} tweets downloaded so far".format(len(alltweets)))

    #verwijder ReTweets
    for tweet in alltweets:
        if 'RT' in tweet.full_text:
            continue
```



## Effect van Elon Musk's Twitter-gebruik

### Code 2: VADER sentiment

```
import pandas as pd
import matplotlib as plt
from datetime import date
!pip install VaderSentiment
!pip install textblob
#data inladen
df1 = pd.read_csv('tweets2.csv', sep=';', encoding='utf-8')
#kolommen invoegen
df1['created_at'].head()
df1['created_at'] = pd.to_datetime(df1['created_at'], unit='ns')
df1['date'] = pd.to_datetime(df1['created_at']).dt.date
df1['date'] = pd.to_datetime(df1['date'])
df1.head()
df1.info()
#data inladen
df2 = pd.read_csv('tweets3_tab.txt', sep='\t')
#kolommen invoegen
df2.sample()
df2['date'] = pd.to_datetime(df2['created_at'], format='%d-%m-%Y')
df2 = df2.drop(['Unnamed: 3', 'Unnamed: 4'], axis=1)
df2.head()
tweets = df1.append(df2, ignore_index=True)
tweets.head()
tweets['date'].value_counts().head(10)
tweets['text'][1]
#fout met b' voor welke tweet verwijderen
tweets['text'][1].replace("b'", "").replace('b"', '')[:-1]
tweets['text'] = tweets['text'].apply(lambda x: x.replace("b'",
 "").replace('b"', '')[:-1])
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
from textblob import TextBlob
#initialiseer sentiment
analyzer= SentimentIntensityAnalyzer()
sentence = tweets['text'][234]
print(sentence)
tweets['compound'] = [analyzer.polarity_scores(v)['compound'] for v in
tweets['text']]
tweets['neg'] = [analyzer.polarity_scores(v)['neg'] for v in
tweets['text']]
tweets['neu'] = [analyzer.polarity_scores(v)['neu'] for v in
tweets['text']]
tweets['pos'] = [analyzer.polarity_scores(v)['pos'] for v in
tweets['text']]
tweets.head()
tweets.to_csv(f'elonmusk_sentiment_{date.today()}.csv')
```

## Effect van Elon Musk's Twitter-gebruik

### Code 3: Word cloud

```
pip install wordcloud
import numpy as np
import pandas as pd
from os import path
from PIL import Image
from wordcloud import WordCloud, STOPWORDS, ImageColorGenerator
import matplotlib.pyplot as plt
%matplotlib inline
#data inladen
df = pd.read_csv('tweets4.csv', sep=';')
df.head()
text = df['text'][1]
df['text'][1]
#opmaak woordenwolk
wordcloud = WordCloud().generate(text)
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis("off")
plt.show()
wordcloud = WordCloud(max_font_size=50, max_words=100,
background_color="white", random_state=42).generate(text)
plt.figure()
plt.imshow(wordcloud, interpolation="bilinear")
plt.axis("off")
plt.show()
all_text = " ".join(x for x in df['text'])
all_text
#stopwoorden invoegen
stopwords = set(STOPWORDS)
stopwords.update(['co', 'https', 'nhttps', 'nhttp', 'x80', 'xa6', 'xe2',
'x80', 'amp'])
#afmetingen woordenwolk
wordcloud_all = WordCloud(stopwords=stopwords, max_font_size=40,
max_words=50, scale=3, random_state=1,
background_color="white").generate(all_text)
plt.imshow(wordcloud_all, interpolation='bilinear')
plt.axis("off")
plt.show()
```

### Appendix IV: Relevante python modules + versies

Met Python 3.8:

```
alabaster==0.7.12
anaconda-client==1.7.2
anaconda-navigator==1.9.12
anaconda-project==0.8.3
argh==0.26.2
asn1crypto==1.3.0
astroid==2.3.3
astropy==4.0
atomicwrites==1.3.0
attrs==19.3.0
autopep8==1.4.4
Babel==2.8.0
backcall==0.1.0
backports.functools-lru-cache==1.6.1
backports.shutil-get-terminal-size==1.0.0
backports.tempfile==1.0
backports.weakref==1.0.post1
bcrypt==3.1.7
beautifulsoup4==4.8.2
bitarray==1.2.1
bkcharts==0.2
bleach==3.1.0
bokeh==1.4.0
boto==2.49.0
Bottleneck==1.3.2
certifi==2019.11.28
cffi==1.14.0
chardet==3.0.4
Click==7.0
cloudpickle==1.3.0
clyent==1.2.2
colorama==0.4.3
comtypes==1.1.7
conda==4.8.2
conda-build==3.18.11
conda-package-handling==1.6.0
conda-verify==3.4.2
contextlib2==0.6.0.post1
cryptography==2.8
cyclер==0.10.0
Cython==0.29.15
cytoolz==0.10.1
dask==2.11.0
decorator==4.4.1
defusedxml==0.6.0
diff-match-patch==20181111
distributed==2.11.0
docutils==0.16
entrypoints==0.3
et-xmlfile==1.0.1
fastcache==1.1.0
filelock==3.0.12
flake8==3.7.9
Flask==1.1.1
fsspec==0.6.2
```

## Effect van Elon Musk's Twitter-gebruik

```
future==0.18.2
gevent==1.4.0
glob2==0.7
greenlet==0.4.15
h5py==2.10.0
HeapDict==1.0.1
html5lib==1.0.1
hypothesis==5.5.4
idna==2.8
imageio==2.6.1
imagesize==1.2.0
importlib-metadata==1.5.0
intervaltree==3.0.2
ipykernel==5.3.0
ipython==7.12.0
ipython-genutils==0.2.0
ipywidgets==7.5.1
isort==4.3.21
itsdangerous==1.1.0
jdcal==1.4.1
jedi==0.14.1
Jinja2==2.11.1
joblib==0.14.1
json5==0.9.1
jsonschema==3.2.0
jupyter==1.0.0
jupyter-client==5.3.4
jupyter-console==6.1.0
jupyter-core==4.6.1
jupyterlab==1.2.6
jupyterlab-server==1.0.6
keyring==21.1.0
kiwisolver==1.1.0
lazy-object-proxy==1.4.3
libarchive-c==2.8
llvmlite==0.31.0
locket==0.2.0
lxml==4.5.0
MarkupSafe==1.1.1
matplotlib==3.1.3
mccabe==0.6.1
menuinst==1.4.16
mistune==0.8.4
mkl-fft==1.0.15
mkl-random==1.1.0
mkl-service==2.3.0
mock==4.0.1
more-itertools==8.2.0
mpmath==1.1.0
msgpack==0.6.1
multipledispatch==0.6.0
navigator-updater==0.2.1
nbconvert==5.6.1
nbformat==5.0.4
networkx==2.4
nltk==3.4.5
nose==1.3.7
notebook==6.0.3
```

## Effect van Elon Musk's Twitter-gebruik

```
numba==0.48.0
numexpr==2.7.1
numpy==1.18.1
numpydoc==0.9.2
oauthlib==3.1.0
olefile==0.46
openpyxl==3.0.3
packaging==20.1
pandas==1.0.1
pandocfilters==1.4.2
paramiko==2.7.1
parso==0.5.2
partd==1.1.0
path==13.1.0
pathlib2==2.3.5
pathtools==0.1.2
patsy==0.5.1
pep8==1.7.1
pexpect==4.8.0
pickleshare==0.7.5
Pillow==7.0.0
pkginfo==1.5.0.1
pluggy==0.13.1
ply==3.11
prometheus-client==0.7.1
prompt-toolkit==3.0.3
psutil==5.6.7
py==1.8.1
pycodestyle==2.5.0
pycosat==0.6.3
pycparser==2.19
pycrypto==2.6.1
pycurl==7.43.0.5
pydocstyle==4.0.1
pyflakes==2.1.1
Pygments==2.5.2
pylint==2.4.4
PyNaCl==1.3.0
pyodbc===4.0.0-unsupported
pyOpenSSL==19.1.0
pyparsing==2.4.6
pyreadline==2.1
pyrsistent==0.15.7
PySocks==1.7.1
pytest==5.3.5
pytest-arraydiff==0.3
pytest-astropy==0.8.0
pytest-astropy-header==0.1.2
pytest-doctestplus==0.5.0
pytest-openfiles==0.4.0
pytest-remotedata==0.3.2
python-dateutil==2.8.1
python-jsonrpc-server==0.3.4
python-language-server==0.31.7
pytz==2019.3
PyWavelets==1.1.1
pywin32==227
pywin32-ctypes==0.2.0
```

## Effect van Elon Musk's Twitter-gebruik

```
pywinpty==0.5.7
PyYAML==5.3
pyzmq==18.1.1
QDarkStyle==2.8
QtAwesome==0.6.1
qtconsole==4.6.0
QtPy==1.9.0
requests==2.22.0
requests-oauthlib==1.3.0
rope==0.16.0
Rtree==0.9.3
ruamel-yaml==0.15.87
scikit-image==0.16.2
scikit-learn==0.22.1
scipy==1.4.1
seaborn==0.10.0
Send2Trash==1.5.0
simplegeneric==0.8.1
singledispatch==3.4.0.3
six==1.14.0
snowballstemmer==2.0.0
sortedcollections==1.1.2
sortedcontainers==2.1.0
soupsieve==1.9.5
Sphinx==2.4.0
sphinxcontrib-applehelp==1.0.1
sphinxcontrib-devhelp==1.0.1
sphinxcontrib-htmlhelp==1.0.2
sphinxcontrib-jsmath==1.0.1
sphinxcontrib-qthelp==1.0.2
sphinxcontrib-serializinghtml==1.1.3
sphinxcontrib-websupport==1.2.0
spyder==4.0.1
spyder-kernels==1.8.1
SQLAlchemy==1.3.13
statsmodels==0.11.0
sympy==1.5.1
tables==3.6.1
tblib==1.6.0
terminado==0.8.3
testpath==0.4.4
textblob==0.15.3
toolz==0.10.0
tornado==6.0.3
tqdm==4.42.1
traitlets==4.3.3
tweepy==3.9.0
ujson==1.35
unicodcsv==0.14.1
urllib3==1.25.8
vaderSentiment==3.3.2
watchdog==0.10.2
wcwidth==0.1.8
webencodings==0.5.1
Werkzeug==1.0.0
widgetsnbextension==3.5.1
win-inet-pton==1.1.0
win-unicode-console==0.5
```

## Effect van Elon Musk's Twitter-gebruik

```
wincertstore==0.2  
wordcloud==1.7.0  
wrapt==1.11.2  
xlrd==1.2.0  
XlsxWriter==1.2.7  
xlwings==0.17.1  
xlwt==1.3.0  
xmltodict==0.12.0  
yapf==0.28.0  
zict==1.0.0  
zipp==2.2.0
```