

Oorzaken en gevolgen van fluctuaties in voedselprijzen in niet-Westerse landen

Janne Spijkervet 10879609
Jardenna Mohazzab 11701145
Jonne Goedhart 11890169
Julius Huizing 11898402

28 juni 2018

Inleiding

De groei van de internationale handel heeft er voor gezorgd dat meer landen toegang hebben tot meer buitenlandse producten. Maar het zorgde ook voor meer afhankelijkheid tussen landen, zoals duidelijk werd tijdens de wereldwijde voedselcrisis van 2007. Van veel breed-geconsumeerde producten stegen de prijzen in korte tijd onvoorspelbaar hoog. Zo steeg de wereldwijde prijs van rijst binnen twee jaar met 217% en de prijs van graan met 136% [1]. Hoewel dit ook in ontwikkelde landen voor veel economische en sociale onrust zorgde, bleken hierdoor vooral de ontwikkelingslanden nog armer te worden [2].

Er is veel onderzoek gedaan naar mogelijke oorzaken van prijsstijgingen, met als doel een herhaling van de voedselcrisis te voorkomen. Zo blijkt het toenemende gebruik van biobrandstoffen de voedselprijzen te verhogen. Immers kunnen gewassen die gebruikt worden voor de productie van brandstof niet meer worden gebruikt voor de productie van voedsel [3]. Ook blijkt dat de olieprijs positief gecorreleerd is aan voedselprijzen, omdat het zowel de energiebron is voor mechanische voedselproductie als voor het transporteren ervan [4]. Bovendien vergroot een stijging van de olieprijs weer de vraag naar biobrandstoffen.

Naast brandstofgebruik en olieprijsen kunnen echter nog meer factoren in verband staan met fluctuaties van voedselprijzen, zoals geografische factoren. Daaruit volgt de volgende deelvraag: *Vertonen landen in dezelfde regio's vergelijkbare prijsverschillen?* Er wordt verondersteld dat landen in dezelfde sub-regio's soortgelijke fluctuaties in voedselprijzen over dezelfde gewassen laten zien. Dit is aannemelijk omdat landen in dezelfde regio's een gelijksoortig klimaat kennen. Dit wordt onderzocht door met k-means landen te classificeren. Verwacht wordt dat het clusteren van landen binnen één regio geen vaste cluster-groepen oplevert.

Daarnaast zijn voedselproducten te categoriseren en bestaan voedselproducten vaak uit andere voedselproducten. Daarom wordt ook geprobeerd de volgende deelvragen te beantwoorden: *Zijn er voedselprijzen die positief of negatief aan elkaar gecorreleerd zijn? Als voedselprijzen aan elkaar gecorreleerd zijn, is dit dan altijd het geval of alleen tijdens bepaalde periodes?* Er wordt verondersteld dat producten uit dezelfde categorieën soortgelijke prijsfluctuaties laten zien, en dat de prijsfluctuatie van een voedselproduct dat uit meerdere ingrediënten bestaat, overeenkomt met de prijsfluctuaties van de individuele ingrediënten waaruit dat product bestaat. Dit wordt onderzocht door verschillende producten te classificeren met k-means. Daarna worden de correlaties tussen producten binnen een cluster met elkaar vergeleken. Verwacht wordt dat producten uit dezelfde categorie samen geclusterd worden en dat correlatie tussen producten binnen deze clusters sterk is. Bovendien wordt verwacht dat de periode daar geen invloed op heeft.

Ook worden voedselprijzen in verschillende valuta gegeven, maar deze valuta zijn niet allemaal even stabiel of gelijk in waarde. Dit leidt tot de volgende deelvraag: *Zijn de wisselkoersen gecorreleerd aan de productprijzen in een land?* Er wordt verondersteld dat er een negatieve lineaire relatie is tussen de wisselkoers (USD per valuta) van een valuta en de prijs van producten. Dit is aannemelijk omdat als de valuta meer waard wordt het goedkoper voor bedrijven is om producten te importeren waardoor de prijzen kunnen dalen. Hier wordt dat onderzocht door de gemiddelde correlatie tussen alle productprijzen en wisselkoersen te bepalen. Er wordt verwacht als de wisselkoers hoger wordt dat de prijs van dat product zal dalen.

Bovendien bestaan er vaak ook grote verschillen in welvaart tussen landen. Dat leidt tot de volgende deelvraag: *Geven prijzen gecorrigeerd op het Gross Domestic Product (GDP) een ander beeld dan prijzen in USD?* Er wordt vanuit gegaan dat het GDP van een land niet gecorreleerd is met productprijzen. Bovendien wordt verondersteld dat corrigeren op GDP een betrouwbaarder beeld geeft van de betaalbaarheid van producten. Immers kan de GDP tussen landen erg verschillen. Hier wordt dat onderzocht door landen met k-means te classificeren op voedselprijzen in USD. Vervolgens worden alle landen opnieuw geclassificeerd, maar dan met prijzen gecorrigeerd op GDP. Verwacht wordt dat dit tot verschillende classificaties leidt.

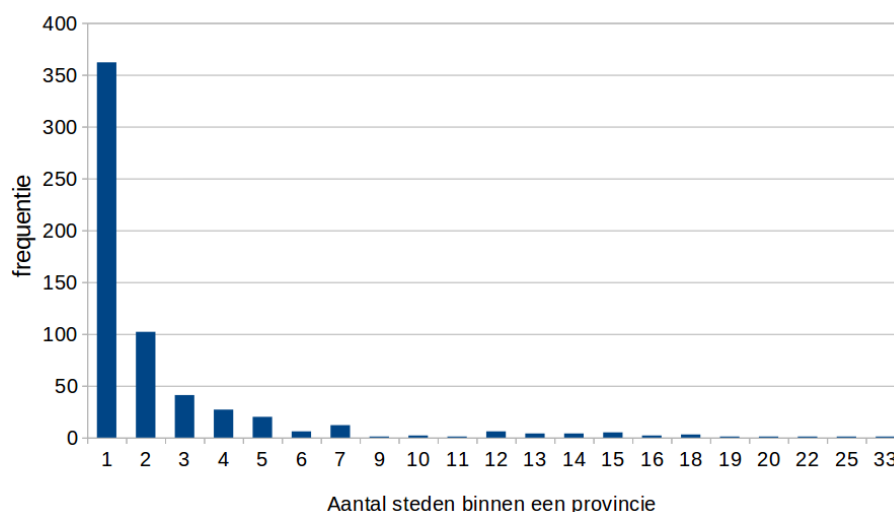
Tot slot wordt de beschikbaarheid van voedsel vaak in verband gebracht met gezondheid, daarom wordt ook geprobeerd de volgende deelvragen te beantwoorden: *Bestaat er een verband tussen de voedselprijzen en het sterftecijfer in een land? Bestaat er een verband tussen vluchtelingen-stromingen en voedselprijzen?* Er wordt vanuit gegaan dat hoge voedselprijzen een negatieve invloed heeft op de algemene gezondheid in een land. Bovendien wordt verwacht dat hoge prijzen een stimulans voor emigratie kunnen zijn. Hier wordt dat onderzocht door voedselprijzen met sterfte- en vluchtelingen cijfers te vergelijken. Er wordt een positieve correlatie verwacht tussen vluchtelingenstromen, sterftecijfers en voedselprijzen.

Om deze deelvragen te kunnen beantwoorden is gebruik gemaakt van de Global Food Prices (GFP) Database van het World Food Programme [5]. In de database zijn voor 76 verschillende landen de maandelijkse voedselprijzen gegeven voor veel geconsumeerde voedselwaren zoals bonen, rijst en olie. Deze database wordt uitgebreid genormaliseerd zodat zowel de voedselprijzen van verschillende producten met elkaar vergeleken kunnen worden als de prijsveranderingen tussen landen en gebieden.

Methode

Pre-processing

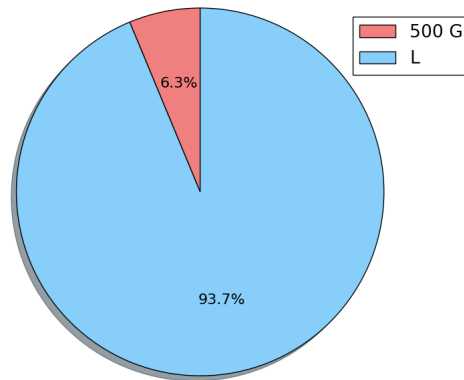
Ten eerste is het aantal kolommen van de dataset verminderd. De kolom met de provincies is verwijderd omdat bij het merendeel van de provincies enkel data was verzameld van één stad, waardoor deze kolom geen extra informatie bevatte dan de kolom met steden zelf (zie figuur 1). Ook waren in de oorspronkelijke dataset de jaren en de maanden gegeven in aparte kolommen. Deze zijn samen-gevoegd tot één kolom.



Figuur 1: Het aantal provincies met hetzelfde aantal gedocumenteerde steden

In de GFP Database was bij veel provincies slechts data van één stad gedocumenteerd.

In de database waren ook veel producten in verschillende eenheden gedocumenteerd. Zo was bijvoorbeeld de melkprijs zowel per 500 gram als per liter gegeven (zie figuur 2). Daarom zijn alle eenheden in de dataset genormaliseerd om het vergelijken van prijzen te vergemakkelijken. Daarvoor is handmatig een *dictionary* aangemaakt met als *keys* alle te converteren eenheden en met als *values* de desbetreffende *conversion rates*. Zo zijn de melkprijzen in liters gedeeld door de *value* van liter in de *dictionary*, waardoor alle melkprijzen nu in KG zijn gegeven. Op deze manier is ervoor gezorgd dat alle producten in een unieke eenheid vermeld staan. De *conversion rates* zijn verkregen uit *matrix conversion rates*[6] en *density conversion rates*[7]. Eenheden waarvoor geen *conversion-rate* gevonden kon worden, zijn de betreffende rijen van verwijderd, zoals bij bijvoorbeeld de eenheden *quartillas* en *goat heads*.



Figuur 2: De verdeling van melk eenheden voor normalisatie

Om voedselprijzen met elkaar te kunnen vergelijken, was het noodzakelijk alle voedselprijzen te normaliseren naar een unieke eenheid.

Bovendien is voor elk land in de database de voedselprijs in verschillende wisselkoersen gedocumenteerd. Om prijzen tussen landen te vergelijken was het daarom noodzakelijk alle prijzen te normaliseren naar één wisselkoers. Hier is ervoor gekozen alle prijzen naar US-Dollar (USD) te normaliseren, omdat de USD één van de meest stabiele wisselkoersen is. Hiervoor is gebruik gemaakt van een historische database van de wisselkoersen [8]. Rijen met valuta's waarvoor geen officiële wisselkoers kon worden gevonden zijn verwijderd. Zo bleek de bijvoorbeeld *Somaliland Shilling* (SOS) geen officieel erkende munteenheid te zijn en was er voor de *Armenian Dram* (AMD) pas wisselkoers-informatie beschikbaar vanaf februari 2003. Zo zijn alle rijen waar de prijs in SOS was gegeven verwijderd, en alle rijen vóór februari 2003 waar de prijs in AMD gegeven was ook.

Om een beter beeld van de betaalbaarheid van producten binnen een land te krijgen, is er een nieuwe maatstaf berekend die rekening houdt met de koopkracht van een land. Hiervoor is gebruik gemaakt van een historische database met GDP's *per capita* [9]. Door telkens de GDP van een land te delen door de prijs van een product in dat land, ontstaat een getal dat aangeeft hoeveel er van dat product in het desbetreffende land gekocht kan worden. Dat getal wordt de 'betaalbaarheid-index' genoemd. Hoe hoger de betaalbaarheid-index van een product in een land, hoe betaalbaarder dat product is voor de inwoners van dat land.

Om voedselprijzen over de tijd met elkaar te vergelijken, is het bovendien noodzakelijk om voor elk product over lange periodes aan onafgebroken voedselprijzen te beschikken. In de dataset zaten echter veel tijds-kloven in de documentatie van voedselprijzen. Kloven van één of twee maanden zijn daarom ingevuld met lineaire regressie. Als er na de toepassing van deze lineaire re-

gressie nog steeds data-reeksen waren van minder dan zes maanden zijn deze reeksen verwijderd. Tenslotte zijn alle producten verwijderd waarvan minder dan één jaar aan data gedocumenteerd is.

Nu alle voedselprijzen over langere periode's met elkaar vergeleken kunnen worden, is voor elke productprijs nog de afgeleide bepaald zodat de daadwerkelijke veranderingen in voedselprijzen vergeleken kunnen worden. En is elk land ingedeeld in sub-regio's zodat ook de voedselprijzen tussen gebieden vergeleken kunnen worden. Hiervoor is gebruik gemaakt van een samengestelde database met alle landen en sub-regio's [10, 11].

Algoritmes

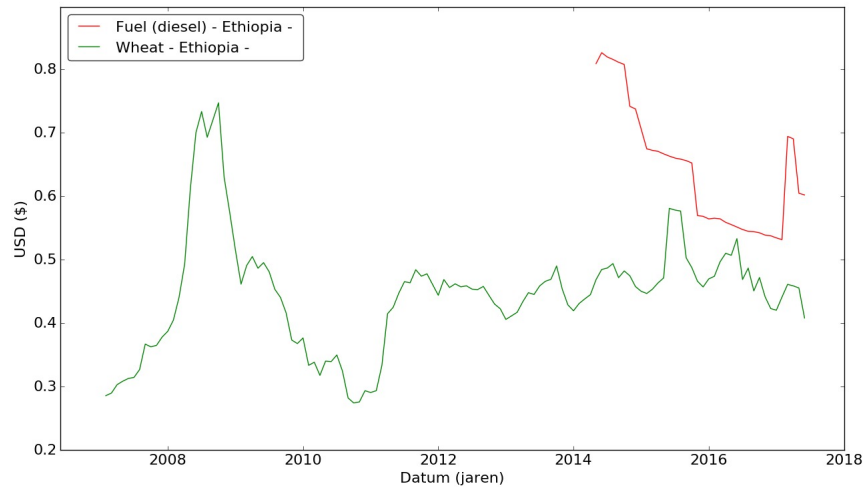
k-means Algoritme

Omdat er verschillende product-prijzen over de tijd met elkaar vergeleken moeten worden, maar niet elk product over dezelfde tijdspan is gedocumenteerd, is het standaard k-means algoritme van *scikit-learn* onbruikbaar. Er ontstaan dan immers *NaN-values* op de tijdstippen waar het ene product wel is gedocumenteerd maar het andere product niet (zie figuur 3). Daarom is er een aangepaste versie van het k-means algoritme geïmplementeerd die deze *NaN-Values* niet meeneemt in de berekeningen. Verder werkt het algoritme op de standaard manier. Dat wil zeggen dat er op willekeurige plekken een door de gebruiker aantal clusters-centra worden geïnitieerd. Vervolgens wordt van elk data-punt de euclidische afstand tot de verschillende centra berekend en worden alle data-punten tot het dichtstbijzijnde cluster-centrum geassocieerd. Dan wordt het gemiddelde van elk cluster berekend en worden de cluster-centra opnieuw geïnitieerd op dat gemiddelde. Tot slot worden alle data-punten opnieuw tot het dichtstbijzijnde cluster-centrum geassocieerd. Dit proces wordt iteratief herhaald tot de totale afstand tussen alle data-punten en hun cluster-centra niet of nauwelijks meer afneemt, met een maximum van 100 iteraties.

Omdat k-means tot lokale-minima kan leiden, wordt in dit onderzoek k-means telkens met dezelfde parameters tien keer herhaald. Verder omdat k-means een *unsupervised learning* algoritme is, wordt het hele proces herhaald met een verschillend aantal groepen. Hierna worden producten die in elke uitkomst van het k-means algoritme hetzelfde worden geassocieerd als cluster beschouwd.

Karl-Pearson correlatie

Om de correlatie tussen verschillende voedselprijzen te bepalen is tenslotte gebruik gemaakt van de Karl-Pearson correlatie coëfficiënt. Dit algoritme geeft een getal terug die de sterkte van de lineaire samenhang tussen twee prijzen aangeeft. Daarbij houdt 1 een perfect lineair verband in en -1 een perfect negatief lineair verband in. Hiermee kan vervolgens worden afgeleid of twee producten afhankelijk zijn van dezelfde factoren of van elkaar.



Figuur 3: Het ontstaan van *NaN-values* bij het vergelijken van prijzen

Wanneer voedselprijzen over verschillende tijdspannen zijn gedocumenteerd, wordt het k-means algoritme van *scikit-learn* onbruikbaar.

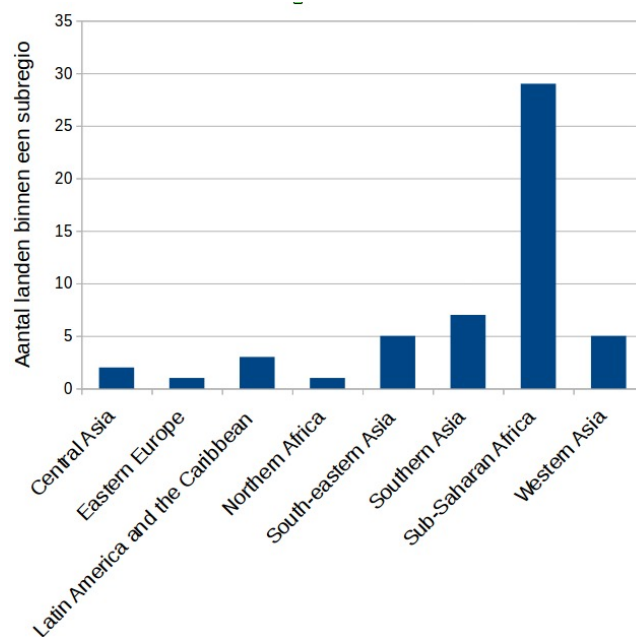
Procedure

Vertonen landen in dezelfde regio's vergelijkbare prijsverschillen?

Om landen binnen een regio goed met elkaar te kunnen vergelijken, moeten er zoveel mogelijk prijsvergelijkingen tussen die landen gemaakt worden. Daarom is gekeken naar de *Sub-Saharan African*, omdat dit de regio is met het grootste aantal gedocumenteerde landen (zie figuur 4). Tussen deze landen zijn vervolgens de kafferkoor (een alternatief voor mais) en de rijst prijzen vergeleken, omdat dit twee van de meest gedocumenteerde producten zijn (zie figuur 5). Vervolgens is geprobeerd de landen te classificeren met k-means, waarna binnen de clusters de correlatiecoëfficiënt tussen de landen is berekend.

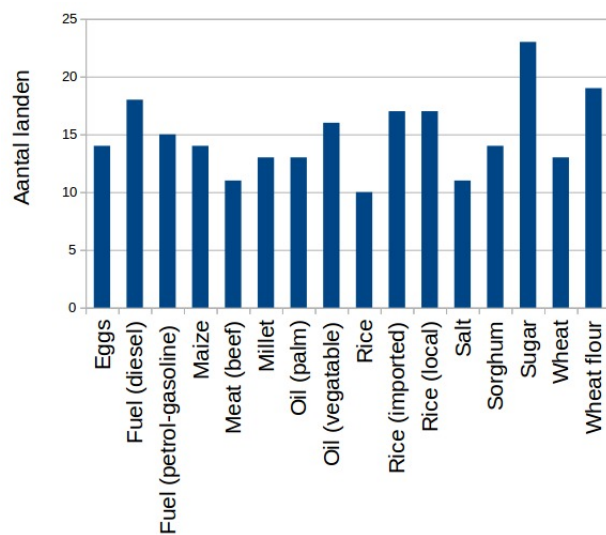
Zijn er voedselprijzen die positief of negatief aan elkaar gecorreleerd zijn? Als voedselprijzen aan elkaar gecorreleerd zijn, is dit dan altijd het geval of alleen tijdens bepaalde periodes?

Om correlaties tussen voedselprijzen te bepalen, zijn er lange onafgebroken periodes aan voedselprijzen nodig. Daarom is er naar de voedselprijzen in Oekraïne gekeken, omdat daar een groot aantal voedselproducten over dezelfde periode (2014-2018) zijn gedocumenteerd. Vervolgens zijn met k-means alle producten geassocieerd, waarna binnen elk cluster de correlatiecoëfficiënt tussen de producten is berekend. Tot slot is gekeken of de classificaties en correlaties veranderen als er naar de periode 2014-2018 wordt gekeken.



Figuur 4: Het aantal landen per sub-regio

De Sub-Sahara blijkt het meeste landen over te hebben na pre-processing.



Figuur 5: Het aantal landen per product

Zijn de wisselkoersen gecorreleerd aan de productprijzen in een land?

Hiervoor is voor elk land de correlatie bepaald tussen de productprijzen en de wisselkoers van dat land. Vervolgens is het gemiddelde van deze correlatie coëfficiënten berekend, waarvan ook de standaard afwijking is berekend.

Geven prijzen gecorrigeerd op het Gross Domestic Product (GDP) een ander beeld dan prijzen in USD?

Eerst is de correlatie tussen GDP en voedselprijzen berekend om te achterhalen hoe betrouwbaar de betaalbaarheid-index is als maatstaf. Als deze correlatie laag is kan de betaalbaarheid-index als zuivere maatstaf worden gebruikt. Vervolgens is naar de graanprijzen gekeken, omdat graan één van de meest-gedocumenteerde producten in de dataset is (zie figuur 5), waardoor er zoveel mogelijk landen met elkaar vergeleken kunnen worden. Eerst zijn de graanprijzen in USD van alle landen in de dataset opgehaald, en zijn de landen op basis daarvan geclassificeerd. Hiervoor is gebruik gemaakt van het k-means algoritme. Vervolgens is de betaalbaarheid-index van graan voor elk land opgehaald en zijn op basis van deze index opnieuw alle landen in clusters ingedeeld. Om te kijken of de prijzen gecorrigeerd met GDP een andere beeld geven wordt er gekeken of clusters van landen uit dezelfde regio leiden tot verschillende classificaties.

Bestaat er een verband tussen de voedselprijzen en het sterftecijfer in een land?

Om een verband tussen voedselprijzen en sterftecijfers te bepalen, is het wenselijk om de cijfers van landen te vergelijken waar sterftecijfers gelijkmatig verlopen. Immers als er een correlatie is tussen de sterftecijfers en voedselprijzen, kunnen de sterftecijfers niet leiden tot sterke fluctuaties in de voedselprijzen. Daarom zijn de voedsel- en sterftecijfers binnen Afghanistan en Niger vergeleken, omdat beide landen gelijkmatige verlopen in sterftecijfers vertonen (zie figuur 12). Vervolgens zijn de grafieken van de voedselprijzen en de sterftecijfers van beide landen vergeleken. Tot slot zijn voor beide landen ook de correlatiecoëfficiënten tussen beide cijfers berekend.

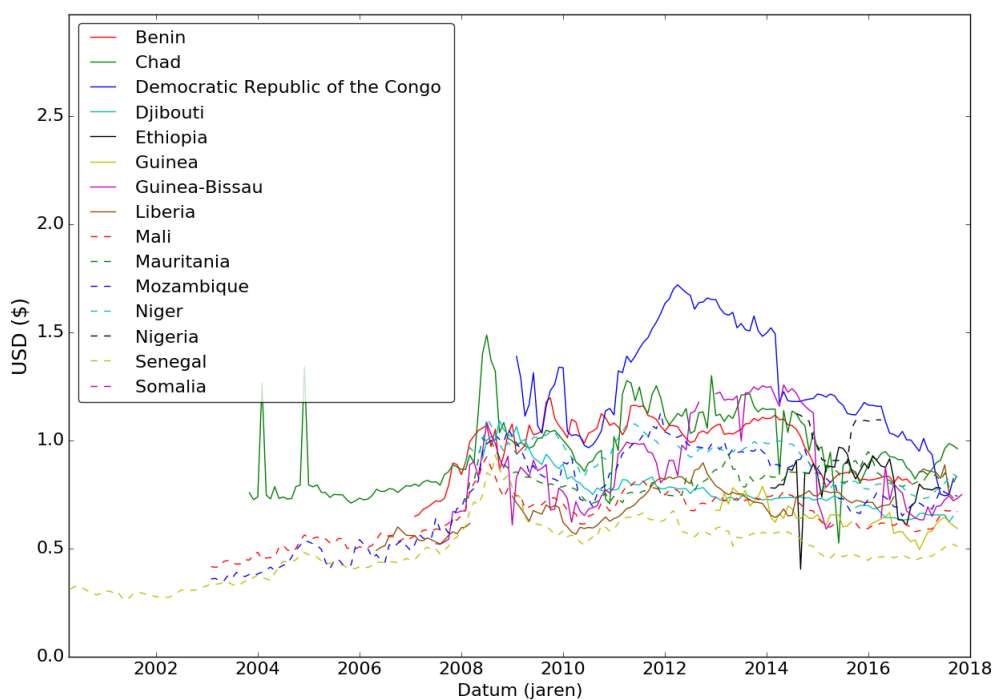
Bestaat er een verband tussen vluchtelingen-stromingen en voedselprijzen?

Om een verband tussen voedselprijzen en vluchtelingen-stromingen te bepalen, is het wenselijk om de cijfers van landen te vergelijken waarbij grote stijgingen van vluchtelingen-cijfers te vinden zijn. Immers kan dan worden gekeken of deze stijgingen samen vallen met stijgingen in voedselprijzen. Daarom is gekeken naar Afghanistan en Georgië, omdat hier grote vluchtelingenstromen te vinden zijn. In beide landen zijn de voedselprijzen en vluchtelingen-cijfers met elkaar vergeleken (zie figuur 15 en 16). Tot slot zijn ook de correlatiecoëfficiënten tussen beide cijfers berekend.

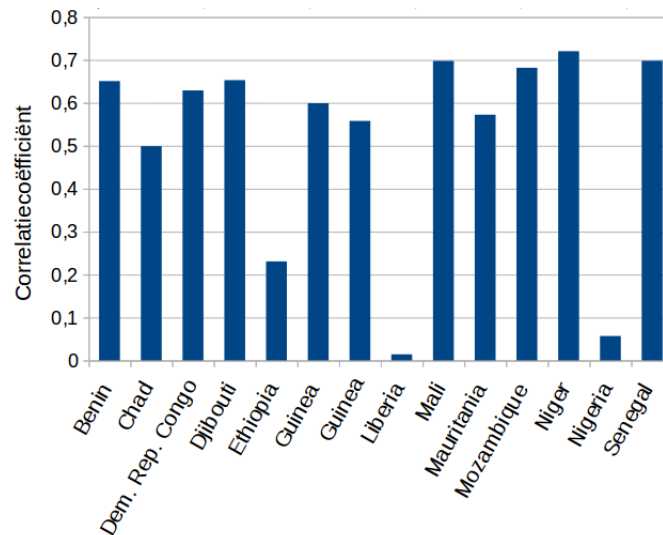
Resultaten

Vertonen landen in dezelfde regio's vergelijkbare prijsverschillen?

Het k-means algoritme clustert de landen telkens in andere groepen. Ook na het herhaald toepassen van het k-means algoritme komen er geen vaste cluster-groepen naar voren. De rijst en kafferkoor prijzen van alle landen lijken gelijksoortige fluctuaties door te maken (zie figuur 6 en 8). Op Ethiopië en Nigeria na, zijn de correlatiecoëfficiënten van de prijzen tussen alle landen aanzienlijk hoog (zie figuur 7 en 9).

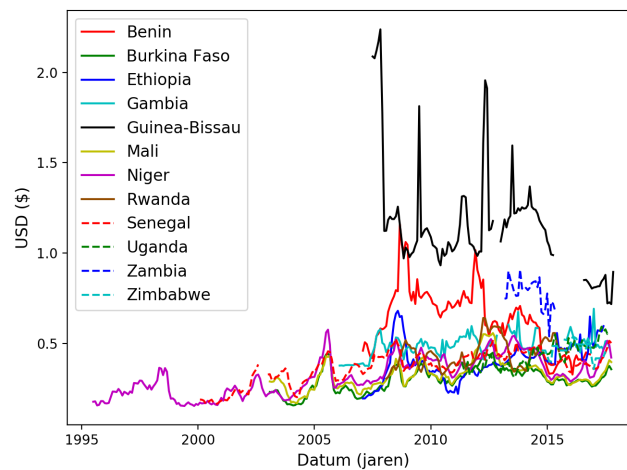


Figuur 6: De prijs fluctuaties van rijst in de landen die behoren tot de *Sub-Saharan Africa*.

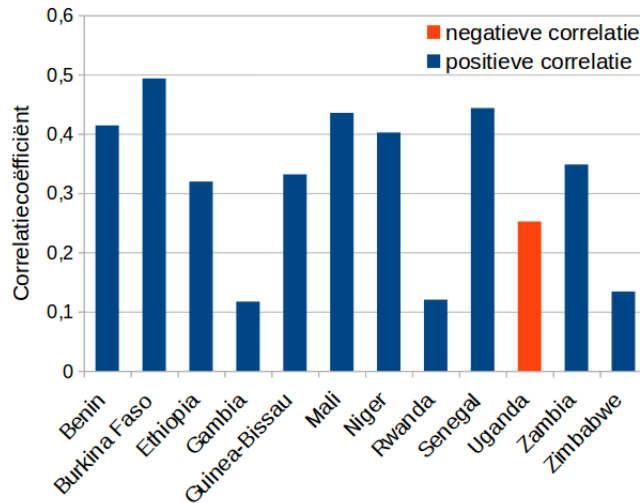


Figuur 7: De gemiddelde onderlinge correlatiecoëfficiënten tussen de rijst-prijzen van alle landen in de *Sub-Saharan Africa*

Op Ethiopië, Liberia en Nigeria na, vertonen alle landen binnen de *Sub-Saharan Africa* sterke prijs correlaties met elkaar.



Figuur 8: De prijs fluctuaties van kafferkoor in de landen die behoren tot Sub-Saharan Africa.



Figuur 9: De gemiddelde onderlinge correlatiecoëfficiënten tussen de kafferkor-prijzen van alle landen in de in de *Sub-Saharan Africa*

De kafferkor-prijzen van de verschillende landen in *Sub-Saharan Africa* vertonen minder correlatie dan de rijst-prijzen.

Zijn er voedselprijzen die positief of negatief aan elkaar gecorreleerd zijn? Als voedselprijzen aan elkaar gecorreleerd zijn, is dit dan altijd het geval of alleen tijdens bepaalde periodes?

Over de periode 2013-2018 komt k-means tot zes clusters (zie tabel 5). Zo worden alle vleesproducten altijd samen in een cluster geplaatst en alle zuivel-producten ook. De prijzen van de producten binnen deze twee clusters correleren sterk (zie tabel 1 en 3), ook als de periode wordt gehalveerd (zie tabel 2 en 4).

Tabel 1: Correlaties tussen zuivel producten in Oekraïne tussen 2014 en 2018

	Butter	Milk	Sour cream
Milk	0.975480		
Sour cream	0.977874	0.996946	
Curd	0.864318	0.930772	0.931195

Tabel 2: Correlaties tussen zuivel producten in Oekraïne tussen 2014 en 2016

	Butter	Milk	Sour cream
Milk	0.989415		
Sour cream	0.994624	0.997019	
Curd	0.985668	0.990619	0.994061

Tabel 3: Correlaties tussen de prijs van vlees producten in Oekraïne tussen 2014 en 2018

	Meat (chicken, whole)	Meat (mixed, sausage)	Meat (pork)
Meat (mixed, sausage)	0.968654		
Meat (pork)	0.943341	0.957977	
Meat (beef)	0.946048	0.962734	0.981646

Tabel 4: Correlaties tussen de prijs van vlees producten in Oekraïne tussen 2014 en 2016

	Meat (chicken, whole)	Meat (mixed, sausage)	Meat (pork)
Meat (mixed, sausage)	0.956601		
Meat (pork)	0.935190	0.957254	
Meat (beef)	0.955861	0.953394	0.969657

Tabel 5: Clusters voor de producten van Oekraïne tussen 2013 en 2018

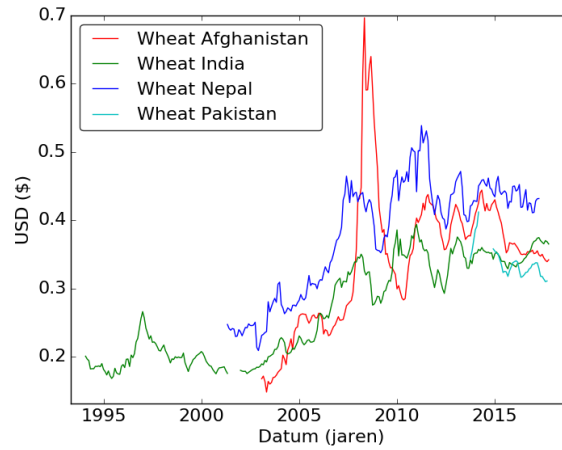
cluster 1	Butter	Curd	Milk	Sour	
cluster 2	Meat (beef)	Meat (chicken, whole)	Meat (mixed, sausage)	Meat (pork)	
cluster 3	Fuel (diesel)	Fuel (petrol-gasoline)			
cluster 4	Cabbage	Carrots	Onions	Potatoes	
cluster 5	Eggs	Pasta	Rice	Sugar	Wheat flour (first)
cluster 6	Bread (rye)	Bread (wheat)	Oil (sunflower)		
overige	Beetroots	Fat (salvo)	Buckwheat grits		

Zijn de wisselkoersen gecorreleerd aan de productprijzen in een land?

De gemiddelde correlatie tussen de productprijzen en de wisselkoers van een land is -0.282759847357, met een standaardafwijking van 0.44391608611.

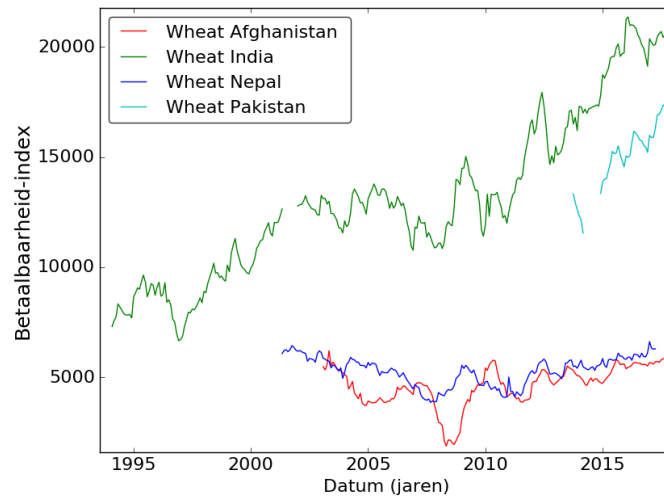
Geven prijzen gecorrigeerd op het Gross Domestic Product (GDP) een ander beeld dan prijzen in USD?

India, Pakistan, Nepal en Afghanistan blijken nabijgelegen landen te zijn die met beide maatstaven verschillend worden geclassificeerd. Met de graanprijzen in USD vertonen India, Pakistan en Nepal gelijksoortige graanprijzen. Deze drie landen worden door k-means dan ook samen geclassificeerd tot het cluster met relatief goedkope graanprijzen. Afghanistan wordt daarentegen geclassificeerd tot het cluster met relatief dure graanprijzen. Echter, wanneer er wordt gekeken naar de betaalbaarheid-index, verandert vooral de betaalbaarheid van graan in Nepal aanzienlijk. Zo wordt Nepal nu samen met Afghanistan tot het cluster met relatief dure graanprijzen geclassificeerd. Bovendien is de gemiddelde correlatie tussen GDP en productprijzen -0.0241052410847, met een standaardafwijking van 0.476439983355.



Figuur 10: De betaalbaarheid van graan in US-Dollars

India, Pakistan en Nepal worden door k-means geclassificeerd tot het cluster met relatief goedkope graanprijzen, terwijl Afghanistan wordt geclassificeerd tot het cluster met relatief dure graanprijzen.

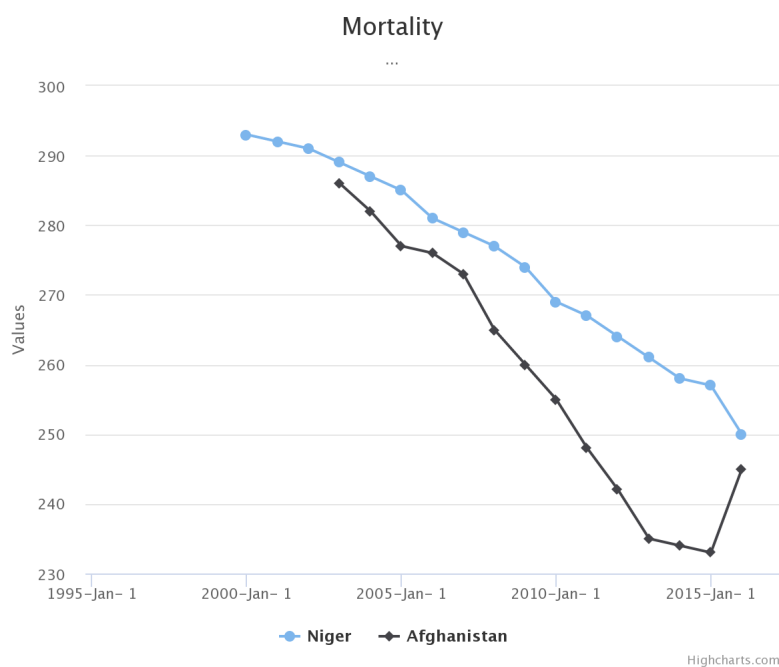


Figuur 11: De betaalbaarheid van graan zoals gegeven door de betaalbaarheid-index

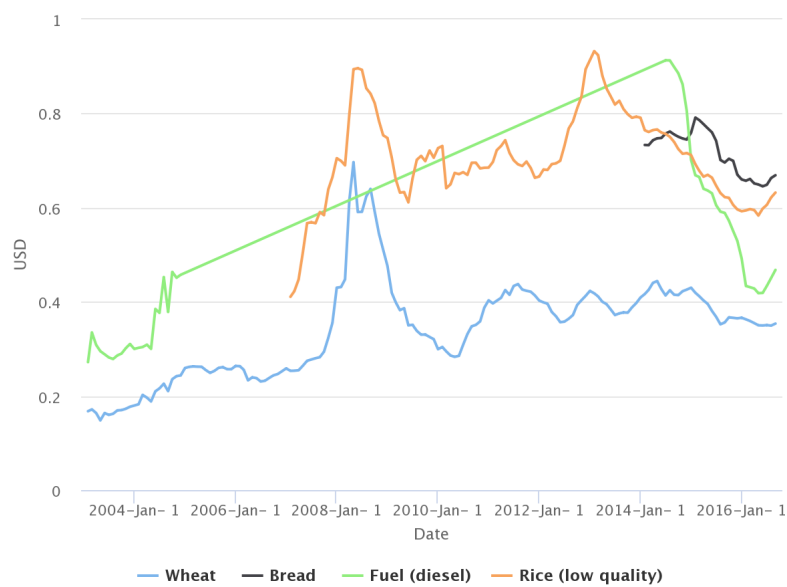
Als er wordt gekeken naar de betaalbaarheid-index, wordt Nepal samen met Afghanistan geclassificeerd tot het cluster met relatief dure graanprijzen.

Bestaat er een verband tussen de voedselprijzen en het sterftecijfer in een land?

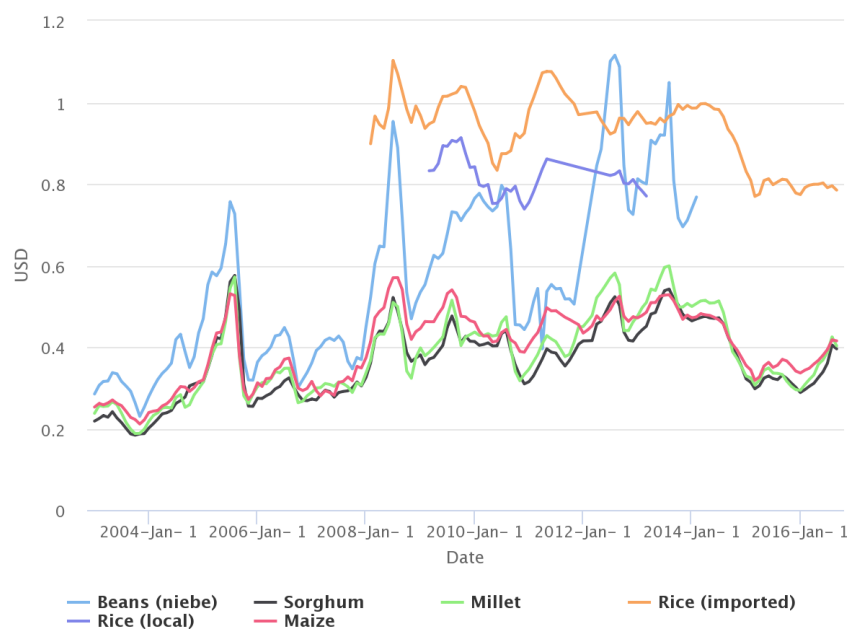
Zoals verwacht verlopen de sterftecijfers van Afghanistan en Niger gelijkmatig (zie figuur 12) in de periode 2000 en 2015. De voedselprijzen in beide landen fluctueren echter onregelmatig tussen 2004 en 2014 (zie figuur 13 en 14). De gemiddelde correlatie van de voedselprijzen in Afghanistan en het sterftecijfer tussen 2004 en 2014 is -0.73. Voor Niger is die correlatie -0.24.



Figuur 12: De sterftecijfers in Afghanistan en Niger in de periode van 2000 tot 2016.



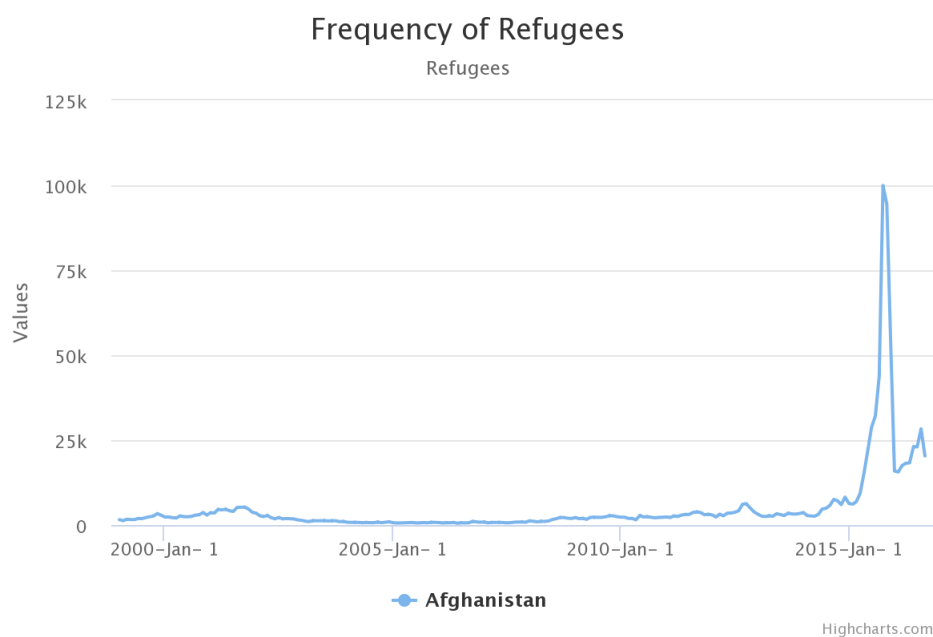
Figuur 13: De voedselprijzen in Afghanistan in de periode van 2004 tot 2016.



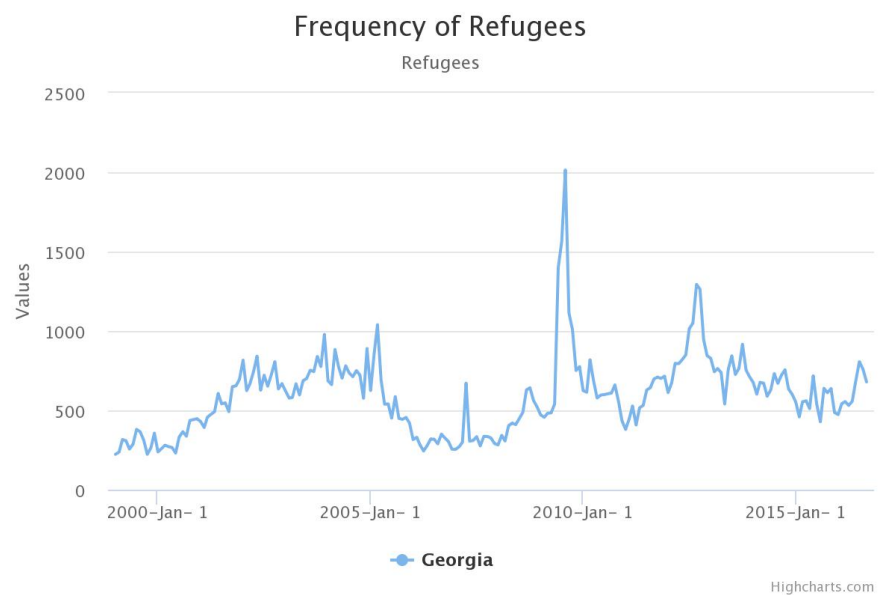
Figuur 14: De voedselprijzen in Niger in de periode van 2004 tot 2016.

Bestaat er een verband tussen vluchtelingen-stromingen en voedselprijzen?

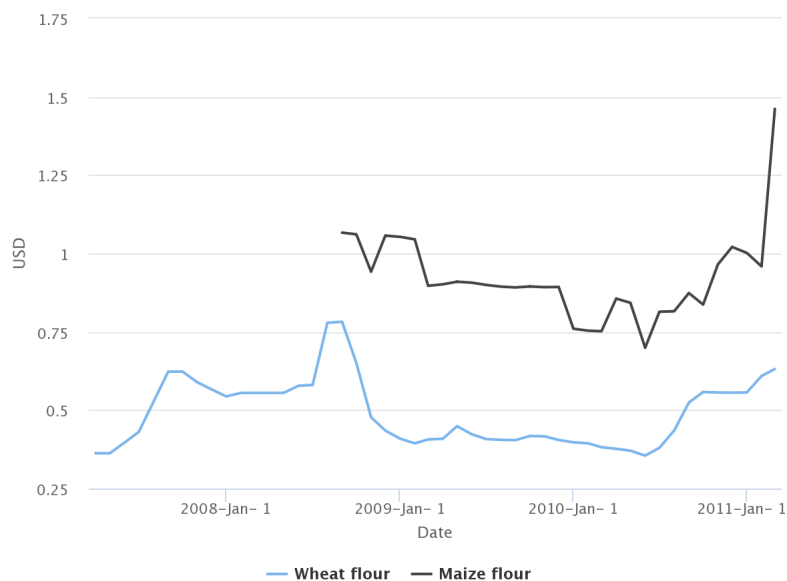
Uit figuur 15 blijkt een grote stijging in het aantal vluchtelingen vanuit Afghanistan rond 2016. In Georgia vindt die grote stijging rond 2010 plaats (zie figuur16). Alle voedselprijzen in Afghanistan laten rond 2016 juist een daling zien (zie figuur 13). Ook de voedselprijzen in Georgia dalen of blijven stabiel rond 2010 (zie figuur 17). De gemiddelde correlatie van de voedselprijzen in Afghanistan en de vluchtelingen tussen 2015 en 2017 is -0.20. Voor Georgia is dat -0.22.



Figuur 15: Het aantal vluchtelingen van Afghanistan in de periode tussen 2000 en 2016



Figuur 16: Het aantal vluchtelingen van Georgië in de periode tussen 2000 en 2016



Figuur 17: De voedselprijzen in Georgië in de periode tussen 2008 en 2011

Discussie

Vertonen landen in dezelfde regio's vergelijkbare prijsverschillen?

Dat het gebruik van k-means niet tot vaste cluster-groepen leidt, kan betekenen dat er sprake is van één cluster. Ook als er naar de grafieken 6 en 8 wordt gekeken kan dit verklaard worden. De prijzen van de verschillende landen vertonen immers allemaal soortgelijke fluctuaties, waardoor k-means waarschijnlijk niet tot consistente classificaties kan komen. Dit is in overeenstemming met de verwachting, omdat het allemaal landen binnen één regio betreft. De relatief lage correlaties van Ethiopië, Liberia en Nigeria betekent mogelijk dat het *outliers* zijn. Dat het k-means algoritme ook deze drie landen samen met de andere landen clustert, desondanks dat ze afwijken, kan verklaard worden doordat k-means geen *outliers* kan detecteren.

Zijn er voedselprijzen die positief of negatief aan elkaar gecorreleerd zijn? Als voedselprijzen aan elkaar gecorreleerd zijn, is dit dan altijd het geval of alleen tijdens bepaalde periodes?

Zoals uit figuur 5 blijkt, worden vleesproducten altijd binnen één groep geclusterd en zuivelproducten ook. Dit is in overeenstemming met de verwachting, omdat dit twee verschillende voedsel-categorieën zijn. De hoge correlaties binnen deze clusters bevestigt ook het vermoeden dat producten uit dezelfde categorie soortgelijke prijsfluctuaties laten zien. Dat het vijfde cluster uit willekeurige producten lijkt te bestaan, komt waarschijnlijk door toevallige correlatie tussen de prijzen van deze producten. Dat *sunflower oil* samen met brood wordt geclassificeerd kan verklaard worden doordat dat mogelijk een veelgebruikt ingrediënt voor brood is, maar ook dit kan door toeval worden verklaard. Tot slot blijkt het kijken naar verschillende periodes vrijwel dezelfde correlaties op te leveren, wat het vermoeden bevestigt dat de periode geen invloed heeft op prijs-correlaties. Om dit met zekerheid te zeggen, zijn echter vergelijkingen over langere periodes nodig.

Zijn de wisselkoersen gecorreleerd aan de productprijzen in een land?

Er blijkt een lichte correlatie (-0.283) te zijn tussen de wisselkoers en productprijzen. De grote standaardafwijking suggereert echter dat deze correlatie per product verschilt. Dit is in overeenstemming met de verwachting, aangezien de wisselkoers vooral invloed heeft op geïmporteerde producten. De grote standaardafwijking valt ook te verklaren doordat een productprijs doorgaans van veel factoren afhangt, waardoor de prijs soms sterker en soms zwakker lijkt te correleren met de wisselkoers. In het vervolg kan daarom onderscheid gemaakt worden tussen geïmporteerde en lokaal-geproduceerde producten, waarbij ook rekening wordt gehouden met andere factoren die van invloed zijn op de prijs van een product, zoals weersomstandigheden en transportkosten.

Geven prijzen gecorrigeerd op het Gross Domestic Product (GDP) een ander beeld dan prijzen in USD?

Dat Nepal eerst tot het goedkopere graanprijs-cluster wordt gerekend, maar aan de hand van de betaalbaarheid-index tot een duurdere cluster, laat zien dat het gebruik van de betaalbaarheid-index een maatstaf is om in overweging te nemen. Waarschijnlijk geeft het vooral bij armere landen een betrouwbaarder beeld van de daadwerkelijke betaalbaarheid van producten, omdat het GDP in deze landen erg laag is. Bovendien blijkt er geen correlatie (-0.024) te zijn tussen GDP en productprijzen. Dit laat zien dat het GDP van een land geen invloed heeft op de productprijzen van dat land, waardoor de betaalbaarheid-index als zuivere maat kan worden gebruikt. Echter zegt de betaalbaarheid van een product nog niet alles over toegang tot dat product. Daarom kan in vervolgonderzoek de betaalbaarheid-index gecombineerd worden met de beschikbaarheid van producten om zo een beeld te geven van de voedselzekerheid in een land.

Bestaat er een verband tussen de voedselprijzen en het sterftecijfer in een land?

Er is geen duidelijk verband te vinden tussen de voedselprijzen en het sterftecijfer in Afghanistan en Niger. Dit kan verklaard worden doordat wanneer de voedselprijzen stijgen, mensen waarschijnlijk over zullen gaan op andere, goedkopere voedselproducten om in leven te blijven. Daarnaast zijn er mogelijk andere factoren die het sterftecijfer sterk beïnvloeden, zoals natuurrampen, waardoor de invloed van voedselprijzen op het sterftecijfer niet is terug te zien in de data. In vervolgonderzoek zou daarom gekeken kunnen worden of de consumptie van relatief goedkope producten stijgt wanneer voedselprijzen stijgen, en of er andere factoren aanwezig zijn die het sterftecijfer kunnen beïnvloeden. De lichte correlaties die gevonden zijn tussen de voedselprijzen en sterftecijfers in beide landen zijn waarschijnlijk het resultaat van toeval.

Bestaat er een verband tussen vluchtelingen-stromingen en voedselprijzen?

Er is geen duidelijk verband te vinden tussen de vluchtelingen-stromingen en de voedselprijzen in Georgia en Afghanistan. Dit kan verklaard worden doordat emigreren vaak lange tijd kost, waardoor stijgingen van voedselprijzen misschien pas later worden gevolgd door toenames van vluchtelingen stromingen. Zo stijgt de prijs van rijst in Afghanistan tussen 2013 en 2014 aanzienlijk (zie figuur 13) en stijgt het aantal vluchtelingen vanuit Afghanistan sterk vanaf 2015 (zie figuur 15). Bovendien kunnen er mogelijk andere factoren zijn die de vluchtelingen stromingen sterk beïnvloeden, zoals oorlogen, waardoor de invloed van voedselprijzen op de vluchtelingen stromingen niet terug te zien is in de data. Tot slot is geen sterke correlatie gevonden tussen beide cijfers.

Referenties

- [1] Stefan Steinberg. “Financial speculators reap profits from global hunger”. In: *Centre for Research on Globalization*. www.globalresearch.ca/printarticle.php (2008).
- [2] Maros Ivanic en Will Martin. “Implications of higher global food prices for poverty in low-income countries”. In: *Agricultural economics* 39.s1 (2008), p. 405–416.
- [3] Donald Mitchell. “A note on rising food prices”. In: (2008).
- [4] Robert Zubrin en Gal Luft. “Food vs. fuel a global myth”. In: *The Chicago Tribune* 6 (2008).
- [5] *WFP, Global Products*. http://vam.wfp.org/sites/data/WFPVAM_FoodPrices_05-12-2017.csv. Accessed: 2018-06-04.
- [6] *Online Conversion Calculators*. <https://metric-calculator.com/>. Accessed: 2018-06-18.
- [7] *Calculate Volume to Weight of common cooking ingredient*. <http://www.webconversiononline.com/weightof.aspx>. Accessed: 2018-06-18.
- [8] *OANDA, Historical Exchange Rates*. <https://www.oanda.com/fx-for-business/historical-rates>. Accessed: 2018-06-12.
- [9] *IMF, International Monetary Fund*. <https://www.imf.org/external/datamapper/NGDPDPC@WEO/OEMDC/ADVEC/WEOWORLD>. Accessed: 2018-06-15.
- [10] *United Nations, UNSD - Methodology*. <https://unstats.un.org/unsd/methodology/m49>. Accessed: 2018-06-08.
- [11] *Wikipedia, ISO 3166-1*. https://en.wikipedia.org/wiki/ISO_3166-1. Accessed: 2018-06-08, Date Published: 2018-06-11.