

Oorzaken en Gevolgen van Voedsel-Prijsstijgingen in Niet-Westerse Landen

Jardenna, Janne, Jonne, Julius

June 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 in 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% [3]. Hoewel dit ook in ontwikkelde landen voor veel economische en sociale onrust zorgde, bleken hierdoor vooral de ontwikkelingslanden nog armer te worden [1].

Daarom is er 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, omdat hier gewassen voor worden gebruikt die [2]. 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 dat 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 die individuele ingrediënten waaruit dat product bestaat. Bovendien wordt er verondersteld dat er een verband bestaat tussen prijzen van producten met overeenkomstige ingrediënten. Ook bij producten die op een gelijke manier worden geproduceerd en afhangen van dezelfde factoren zal hoogstwaarschijnlijk een correlatie te vinden zijn. Hier wordt dat onderzocht door de prijzen van zuivelproducten met elkaar te vergelijken, en door de prijzen van verschillende soorten vlees met elkaar te vergelijken. Er wordt verwacht dat de prijzen van zuivelproducten sterk correleren. Deze producten hebben immers allemaal een overeenkomstig bestandsdeel, namelijk melk. De prijs van melk heeft dus waarschijnlijk invloed op de prijzen van al deze producten. Wel wordt verwacht dat deze correlatie niet altijd even sterk is, omdat de beschikbaarheid van andere ingrediënten ook van invloed kan zijn op de prijs. Bovendien wordt verwacht dat de prijzen van verschillende soorten vlees dezelfde veranderingen doorgaan.

Bovendien bestaan er vaak ook grote verschillen in welvaart tussen landen. Dat leidt tot de volgende deelvraag:

- Kunnen op US-Dollar genormaliseerde prijzen worden gebruikt om de betaalbaarheid van producten te bepalen, of moet er ook gecorrigeerd worden op het Gross Domestic Product (GDP) van een land?

Er wordt vanuit gegaan dat corrigeren GDP een betrouwbaarder beeld geeft. Immers kan de GDP tussen landen dermate verschillen dat landen die qua prijsfluctuatie in US-dollars gelijk opgaan, na correctie op GDP geheel andere trends laten zien. Hier wordt dat onderzocht door van verschillende landen de voedselprijzen in US-Dollar te bepalen, en deze landen met K-means te laten clusteren tot twee groepen met relatief goedkope en relatief dure rijst-prijzen. Vervolgens is dit voor alle landen herhaald, maar dan met prijzen gecorrigeerd op GDP. Verwacht wordt dat sommigen landen met een laag GDP die met de eerste methode nog tot het goedkope cluster worden gerekend, met de tweede methode in het duurdere cluster worden ingedeeld.

Tot slot wordt de beschikbaarheid van voedsel ook vaak in verband gebracht met gezondheid, daarom worden 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 hebben op de gezondheid, en dat hoge voedselprijzen het aantrekkelijker maakt voor mensen om naar een ander land te vluchten. Hier wordt dat onderzocht door

voedselprijzen met sterftecijfers en vluchtelingen-cijfers te vergelijken. Verwacht wordt dat er een positief verband bestaat tussen de vluchtelingenstromen, sterftecijfers voedselprijzen.

Om deze deelvragen te kunnen beantwoorden is gebruikt gemaakt van de Global Food Prices Database (GFPD) van het World Food Programme (<http://www1.wfp.org/>). In de database zijn voor 76 verschillende landen de maandelijkse voedselprijzen gegeven voor veel geconsumeerde voedsel waren zoals bonen, rijst en olie. Deze database is uitgebreid genormaliseerd zodat zowel de voedselprijzen van verschillende producten met elkaar vergeleken konden worden als de prijsveranderingen tussen landen en gebieden.

Methode

Pre-processing

Ten eerste zijn het 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. Ook waren in de oorspronkelijke dataset de jaren en de maanden gegeven in aparte kolommen. Deze zijn samengevoegd tot één kolom.

Daarna zijn alle eenheden genormaliseerd om het vergelijken van prijsverschillen tussen landen te vergemakkelijken. Daarvoor zijn alle gewichten naar kilogram (KG) en alle volumes naar liter (L) omgezet. Hiervoor is handmatig een *dictionary* aangemaakt met als *keys* alle te converteren eenheden en met als *values* de desbetreffende *conversion rates*. Zo was de melkprijs zowel per liter als per *gallon* gegeven. Om dit te normaliseren zijn de prijzen van *gallon* gedeeld door de *value* van *gallon* in de *dictionary*, namelijk 3.78541178. Om de *conversion rates* te achterhalen is gebruik gemaakt van een online *metric converter* (<https://metric-calculator.com>). Eenheden waarvoor geen vaste *conversion-rate* gevonden kon worden zijn de betreffende rijen verwijderd, zoals bij bijvoorbeeld de eenheden *quartillas* en *goat heads*.

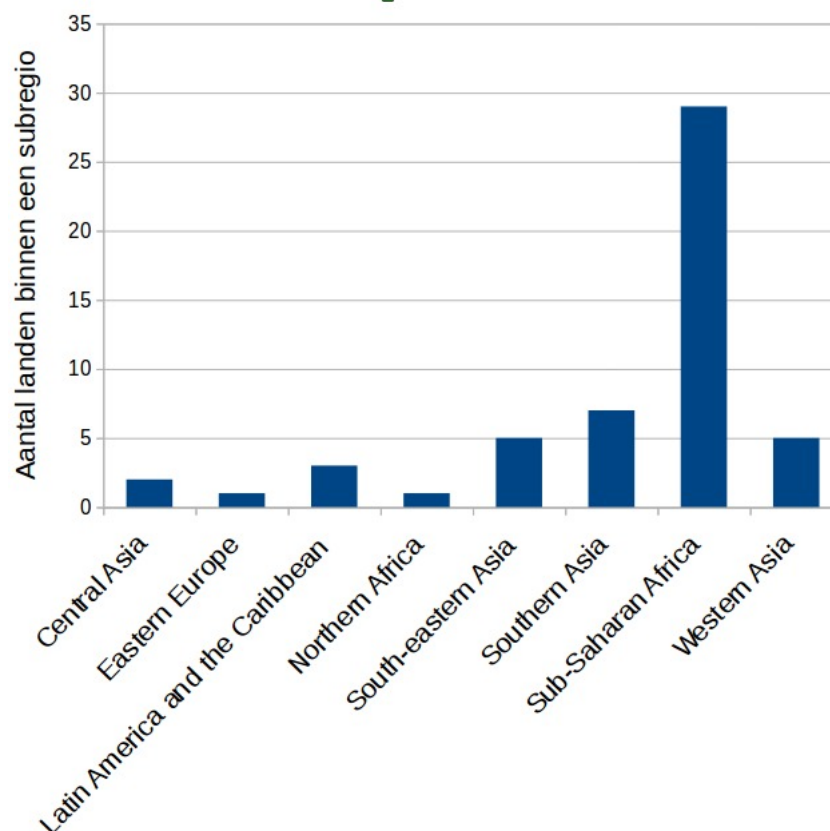
Ook zijn alle wisselkoersen en de bijbehorende prijzen genormaliseerd naar *US Dollars* (USD). Hiervoor is gebruik gemaakt van een historische database van wisselkoersen (<https://www.oanda.com/fx-for-business/historical-rates>). Rij met munteenheden waarvoor geen officiële wisselkoers kon worden gevonden zijn verwijderd. Zo bleek de *Somaliland Shilling* (SOS) geen officieel erkende munteenheid te zijn en was er voor de *Armenian Dram* (AMD) pas wisselkoersinformatie beschikbaar vanaf februari 2003. Alle rijen waar de prijs in SOS was gegeven zijn dus verwijderd, en alle rijen waar de prijs in AMD gegeven was vóór februari 2003 ook.

Omdat nu alle prijzen in USD waren gegeven was het ook mogelijk prijzen per land te normaliseren op *gross domestic product* (GDP). Hiervoor is gebruik gemaakt van een historische database met GDP's *per capita* wereldwijd (<http://www.imf.org>). Door telkens de GDP's van landen door de prijzen van producten in dat land te delen, ontstaat een maat die aangeeft hoeveel er van dat product in het betreffende land gekocht kan worden. Dat getal is de 'betaalbaarheid-index' genoemd. Hoe hoger de betaalbaarheid-index van een product in een land, hoe betaalbaarder dat product voor de inwoners van dat land is.

Tot slot is voor elke productprijs de afgeleide bepaald zodat de daadwerkelijke veranderingen in voedselprijzen tussen steden en landen vergeleken konden worden, en is elk land ingedeeld in sub-regio's zodat ook de voedselprijzen tus-

sen gebieden vergeleken kunnen worden. Hiervoor is gebruik gemaakt van een samengestelde database met alle landen en sub-regio's (<https://unstats.un.org>).

Exploratory Data Analysis

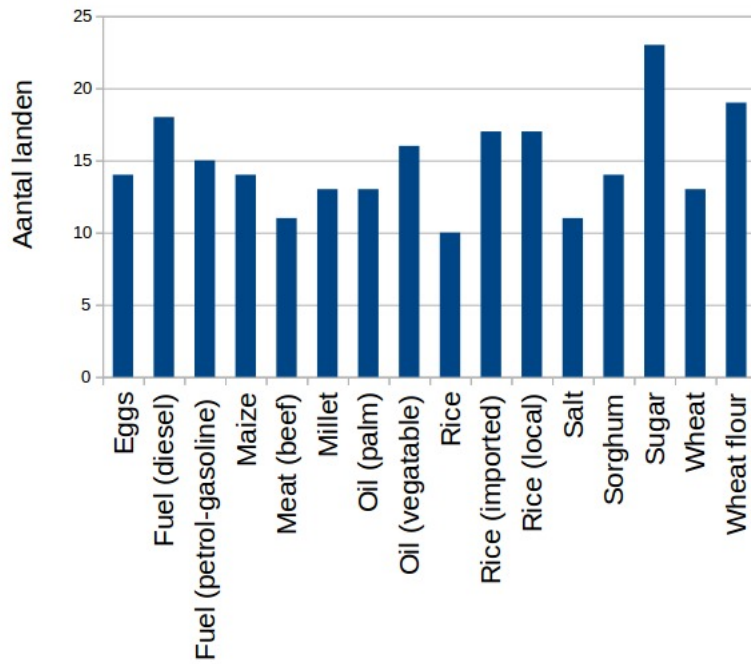


Figuur 1: het aantal landen per sub-regio

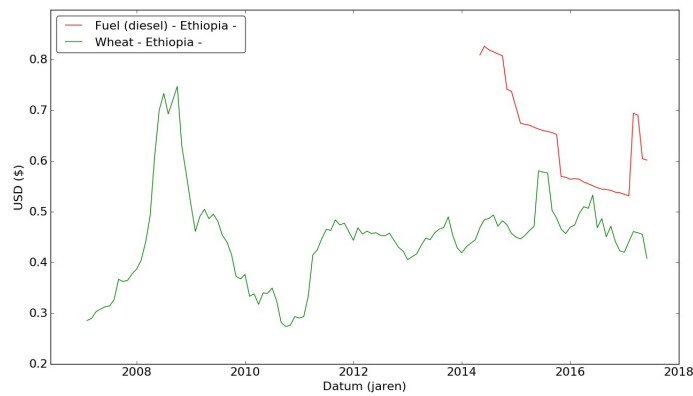
De Sub-Sahara blijkt het meeste landen over te hebben na pre-processing.

lineaire regressie

Omdat het uiteindelijke doel het vergelijken van verschillende voedsel-prijs-grafieken was, en er hiervoor onafgebroken grafieken nodig zijn, zijn alle producten waarvan er meer dan twee opeenvolgende maanden aan productprijzen misten verwijderd. Gaps van één of twee jaar zijn ingevuld met behulp van lineaire regressie, waarbij alleen de eerste productprijs vóór en de eerste productprijs ná de gap als referentiepunten zijn gebruikt.



Figuur 2: het aantal landen per product



Figuur 3:

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

K-Means Algoritme

Omdat er verschillende product-prijzen over de tijd met elkaar vergeleken moesten worden, maar niet elk product over dezelfde tijdspanne was gedocumenteerd,

was 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 X). 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 's 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. Vervolgens wordt het gemiddelde van elk cluster berekend en worden de cluster-centra opnieuw geïnitieerd op dat gemiddelde. Vervolgens 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 dit tot lokale-minima kan leiden, wordt dit nog eens tien keer gedaan. Uiteindelijk wordt de classificatie gekozen met de kortste totale afstand van die tien pogingen.

Data-Analyse

verband olie - en voedselprijzen

Om het verband tussen de olieprijs en de voedselprijzen in Oost-Europa na te gaan is de correlatiecoëfficiënt berekend tussen fluctuaties van de olieprijs en de prijs van 25 verschillende voedselproducten in Oost-Europa, zoals brood melk en eieren. Daarnaast is zowel de gradiënt van de olieprijs als die van de broodprijs bepaald voor de periode 2013-2018. In Zuid-Azië is de gradiënt van de olieprijs met de gradiënt van de thee-prijs vergeleken.

Vervolgens zijn binnen beide werelddelen de gradiënten met elkaar vergeleken en is de correlatiecoëfficiënt berekend met behulp van... Tot slot is voor Oost-Europa de correlatiecoëfficiënt tussen de olieprijs en alle andere voedselproducten berekend.

Vertonen landen in dezelfde regio's vergelijkbare prijsverschillen?

Van de regio Sub-Saharan Africa is veruit de meeste data gedocumenteerd was van verschillende landen (zie figuur 1). Bij de andere regio's was weinig data beschikbaar van overeenkomstige producten tijdens eenzelfde periode, waardoor het vergelijken lastig mogelijk was. Vandaar dat deze regio gekozen is om te onderzoeken voor deze deelvraag. De landen beschikbaar in de dataset die behoren tot de Sub-Saharan Africa zijn: Benin, Chad, Djibouti, Ethiopia, Guinea, Guinea-Bissau, Liberia, Mali, Mauritania, Mozambique, Niger, Nigeria, Senegal. De producten waar naar wordt gekeken zijn rijst en kafferkoor (een Afrikaans alternatief voor mais) omdat er hierover veel data beschikbaar was van landen in de Sub-Saharan. Er is gekeken welke landen bij elkaar geclusterd worden met behulp van het K-means algoritme. Ook zijn de correlatiecoëfficiënten berekend tussen de landen.

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 Oekraïne tussen 2013 en 2018 was data beschikbaar van veel soorten vlees, vandaar dat dit land is gekozen om te onderzoeken voor deze deelvraag. Eerst zijn de genormaliseerde prijzen van alle producten in Oekraïne opgehaald en met behulp van bokeh geplot om een overzicht te krijgen van de relatieve veranderingen.

Met behulp van K-means algoritme is er gekeken welke producten in Oekraïne bij elkaar geclusterd worden. Ook is de correlatie coëfficiënt van Pearson tussen de producten berekend.

Ook was er in Oekraïne veel data beschikbaar over zuivelproducten. Om de mate van correlatie voor deze producten te bepalen is wederom de correlatie coëfficiënt berekend en het K-means algoritme toegepast.

GRAFIEK VAN MELKPRODUCTEN

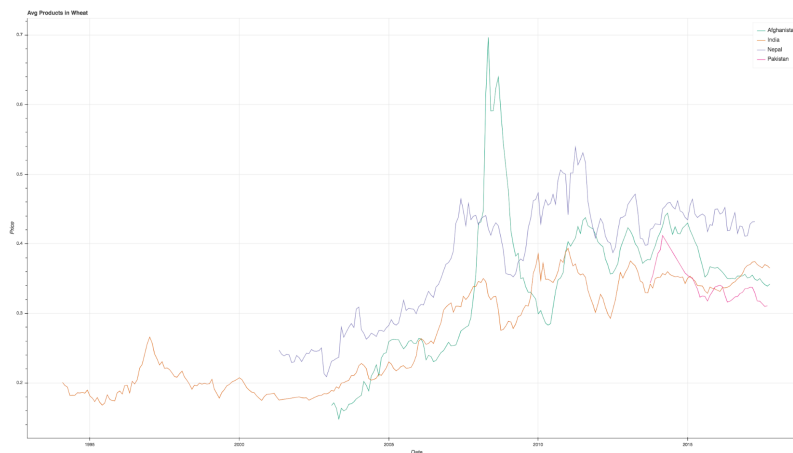
Kunnen op US-Dollar genormaliseerde prijzen worden gebruikt om de betaalbaarheid van producten te bepalen, of moet er ook gecorrigeerd worden op het Gross Domestic Product (GDP) van een land?

Hiervoor zijn alleen naar de graanprijzen gekeken, omdat graan één van de meest-gedocumenteerde producten in de dataset was (zie figuur X), waardoor er zoveel mogelijk landen met elkaar vergeleken konden worden. Eerst zijn de genormaliseerde graanprijzen van alle landen in de dataset opgehaald, en zijn de landen op basis van deze prijs geclassificeerd tot twee groepen. Hiervoor is gebruik gemaakt van het K-means algoritme, waarbij telkens twee clusters werden geïnitieerd. Vervolgens is de betaalbaarheid-index van graan voor elk land opgehaald, en zijn op basis van deze index opnieuw alle landen in twee groepen ingedeeld. Vervolgens is gekeken of prijs in US-dollars en de betaalbaarheid index tot verschillende classificaties leidden.

Resultaten

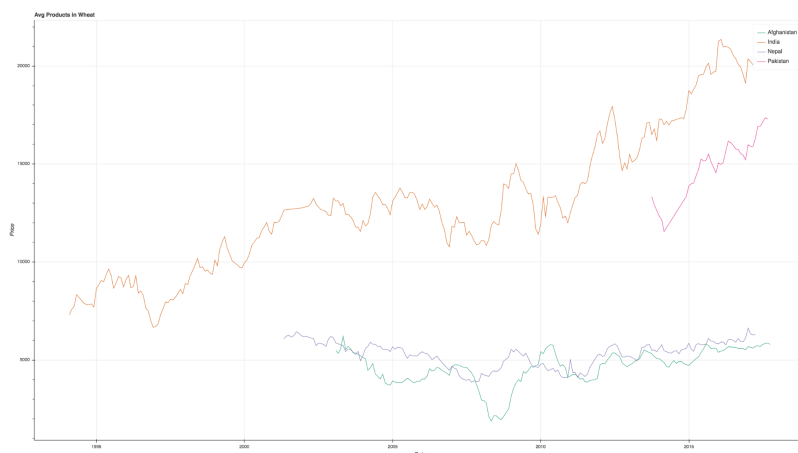
Kunnen op US-Dollar genormaliseerde prijzen worden gebruikt om de betaalbaarheid van producten te bepalen, of moet er ook gecorrigeerd worden op het Gross Domestic Product (GDP) van een land?

Er wordt gekeken naar de verschillende classificaties van Afghanistan, India, Pakistan en Nepal gekeken, omdat de verschillende classificaties van deze landen representatief waren voor de verschillende classificaties van alle landen. Met de graanprijs genormaliseerd op US-Dollars, 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, terwijl Afghanistan wordt geclassificeerd tot het cluster waar de prijzen hoger liggen. Echter, wanneer er wordt gekeken naar de betaalbaarheid-index, verandert vooral de graanprijs van Nepal aanzienlijk. Zo wordt Afghanistan nu samen met Nepal tot het cluster met relatief dure graanprijzen geclassificeerd.



Figuur 4: 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 5: De betaalbaarheid van graan zoals gegeven door de betaalbaarheid-index

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

Tabel 1: Correlaties tussen zuivel producten in Oekraïne tussen 2013 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 de prijs van vlees producten in Oekraïne

	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 3: 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
cluster 6	Bread (rye)	Bread (wheat)	Oil (sunflower)	Wheat flour (first)
overige	Beetroots	Fat (salo)	Buckwheat grits	

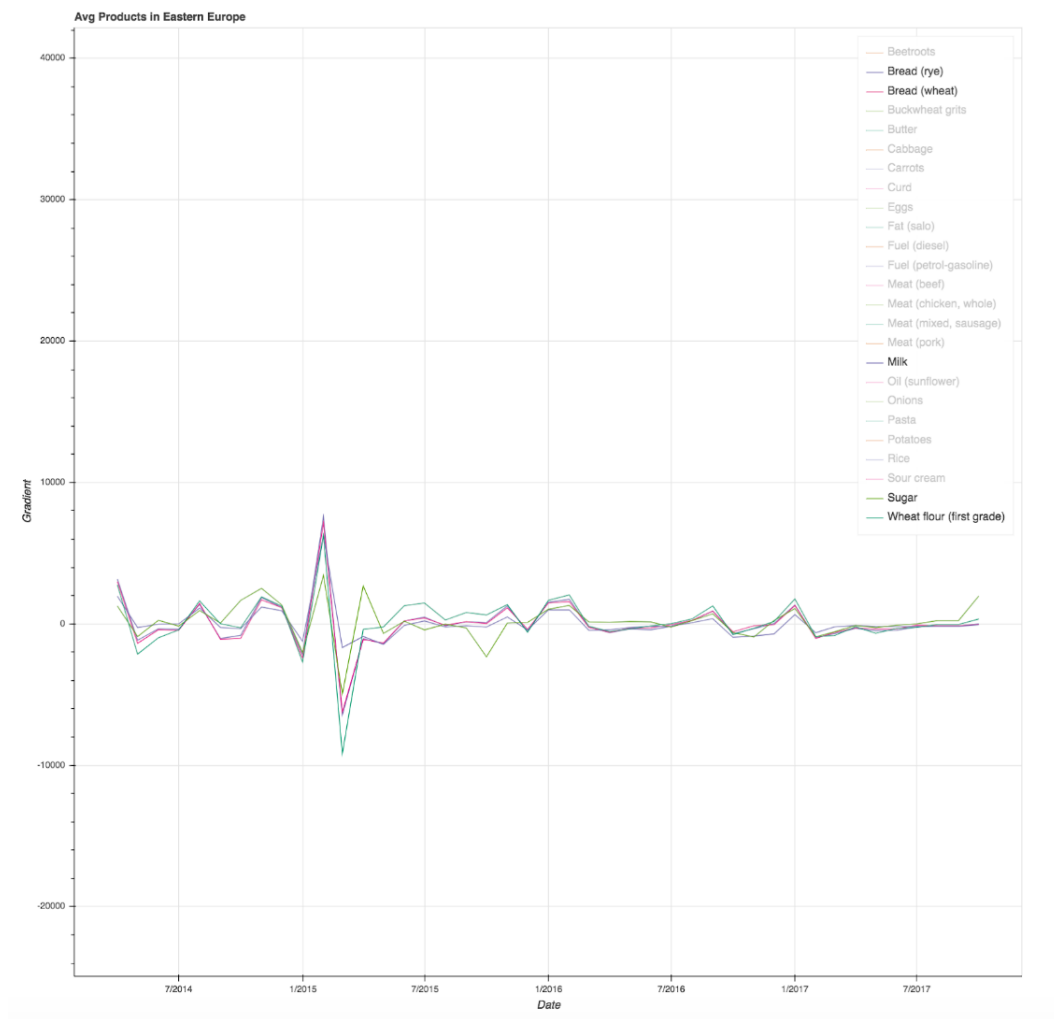
Zijn er voedselprijzen die positief of negatief aan elkaar gecorreleerd zijn? Als voedselprijzen aan elkaar gecorreleerd zijn, is dit dan altijd het geval of alleen tijdens bepaalde periodes?

Vertonen landen in dezelfde regio's vergelijkbare prijsverschillen?

Discussie

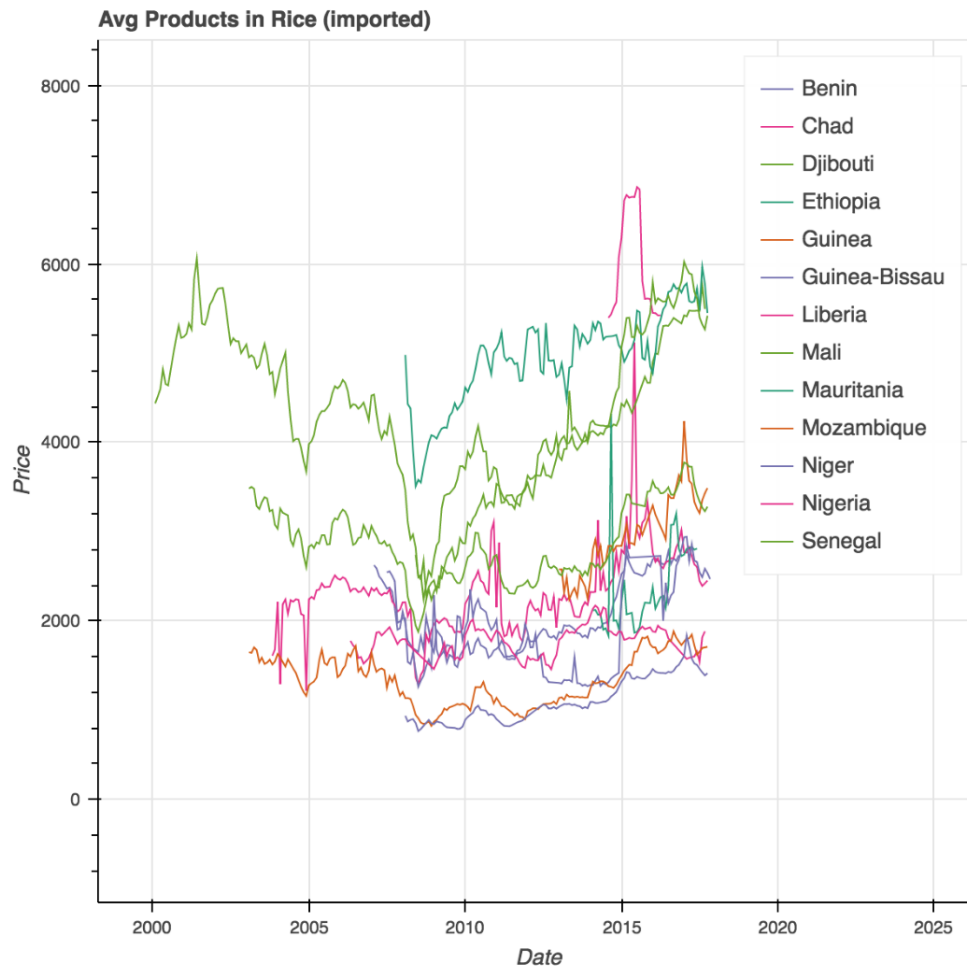
Kunnen op US-Dollar genormaliseerde prijzen worden gebruikt om de betaalbaarheid van producten te bepalen, of moet er ook gecorrigeerd worden op het Gross Domestic Product (GDP) van een land?

Het gebruik van US-Dollar prijzen en de betaalbaarheid-index leiden duidelijk tot andere classificaties. Het Nepal eerst tot het goedkopere graanprijs cluster



Figuur 6: De afnames en toenames in de prijs van de ingrediënten van brood in Oekraïne geplot over een periode van ongeveer 2013 tot 2018.

wordt gerekend gerekend maar aan de hand van de betaalbaarheid-index tot het 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.

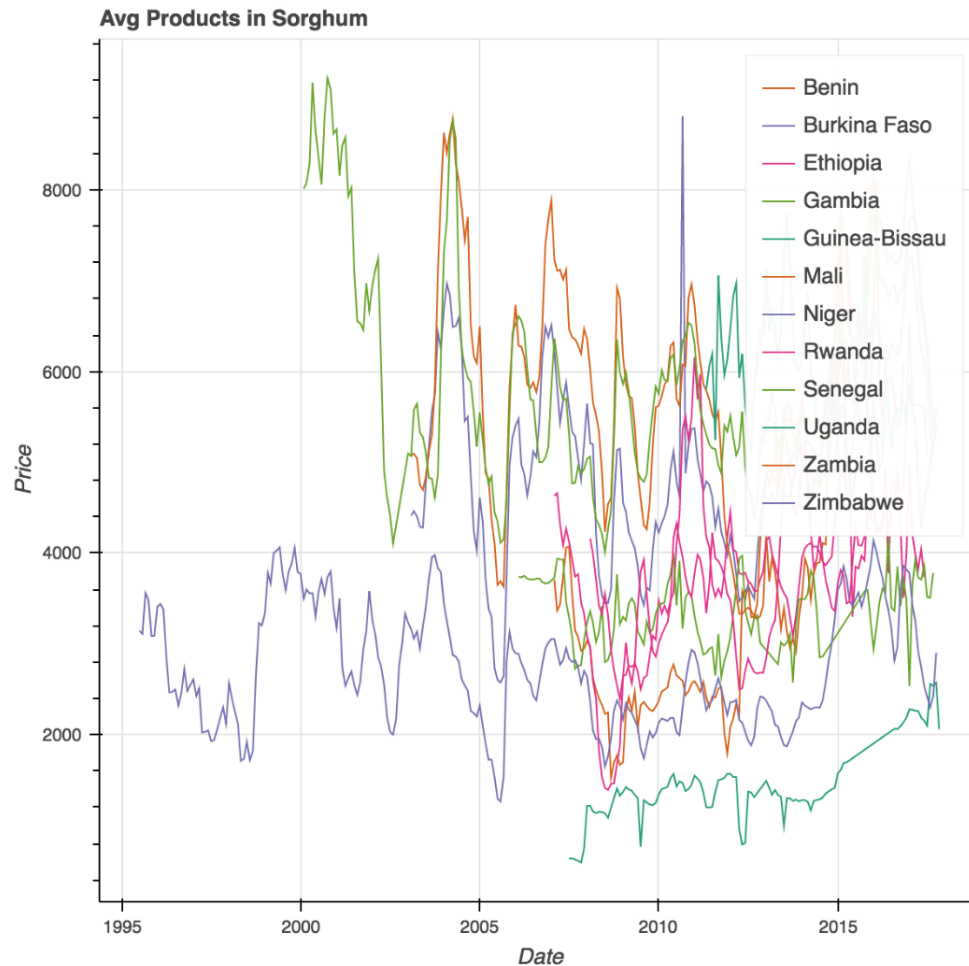


Figuur 7: De afnames en toenames in de prijs van rijst van de landen die behoren tot Sub-Saharan Africa.

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?

Tabel 2 geeft de hoge correlaties weer tussen de verschillende vleessoorten in Oekraïne.

Zoals te zien is in Tabel 1 is de correlatie tussen de zuivel producten in Oekraïne tussen 2013 en 2018 zeer hoog. Bovendien worden deze producten bij elkaar geclusterd door K-means (zie tabel 3). Dit duidt dus op een positieve correlatie tussen de voedselprijzen van zuivelproducten (in Oekraïne). Deze



Figuur 8: De afnames en toenames in de prijs van sorghum van de landen die behoren tot Sub-Saharan Africa.

correlatie was vooral te vinden in de periode tussen 2013 en 2018.

Vertonen landen in dezelfde regio's vergelijkbare prijsverschillen?

K-means geeft als resultaat dat er een verband is tussen vrijwel alle prijzen van rijst van de landen in Sub-Saharan Africa omdat er niet sprake is van consistente clustering. Dat wil zeggen dat het niet mogelijk is om meerdere groepen te vormen omdat de landen als het ware één grote groep vormen. De lineaire correlatie tussen Ethiopia en Nigeria met de andere landen is aanzienlijk laag. Dit is niet gevonden met K-means omdat dit algoritme geen *outliers* kan

detecteren of het kan zijn dat de relatie is niet lineair is. De rijst prijzen van andere landen in Sub-Saharan Africa correleren wel sterk.

Bij het bekijken van de prijzen van kafferkoor valt op dat er een subgroep ontstaat van Burkino-Faso, Mali en Niger tegenover de rest van de landen in *Sub-Saharan Africa*. Met behulp van de correlatie coëfficiënten valt op dat er een hoge correlatie is tussen de prijzen van kafferkoor in landen Mali en Niger, Mali en Burkina Faso, en Niger en Burkina Faso. Dit laat zien dat landen in dezelfde regio's vergelijkbare prijsverschillen vertonen.

Referenties

- [1] 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.
- [2] Donald Mitchell. “A note on rising food prices”. In: (2008).
- [3] Stefan Steinberg. “Financial speculators reap profits from global hunger”. In: *Centre for Research on Globalization*. www.globalresearch.ca/printarticle.php (2008).
- [4] Robert Zubrin en Gal Luft. “Food vs. fuel a global myth”. In: *The Chicago Tribune* 6 (2008).