

'Process Book Data Visualisatie Groep 10'

Tijmen van Etten, Thijs Kuipers, Enno Kuyt & Bart de Rooij

June 2018

Week 1

Deel 1:

Wij zijn begonnen met het ontdekken van de dataset. We hebben uitgezocht hoe compleet de dataset was en besproken wat wij met dit project wilde bereiken. Iedereen is eerst zelf begonnen met werken aan de dataset, om te ontdekken wat de mogelijkheden waren. Er werd een drive bestand aangemaakt waarin TODO lijsten geplaatst konden worden, evenals planningen en voortgang voor het Process Book en de Meeting Protocols.

Deel 2:

We hebben besloten dat we ons specifiek gaan focussen op het midden-oosten. We hadden hier meerdere redenen voor. Zo bevat de dataset relatief veel landen uit deze regio, en vonden wij dit alle vier een interessant gebied. Daarnaast is er de afgelopen jaren veel gebeurd in deze regio, waardoor er veel informatie over te vinden is in de media. Ook kunnen we deze vaak chaotische en opvallende gebeurtenissen, zoals aanslagen of een coup, relatief makkelijk relateren aan afwijkende voedselprijzen.

Week 2

Deel 1:

Na veel verkennend werk met de dataset zijn wij tot de conclusie gekomen dat er voor een aantal van onze zelfbedachte deelvragen, niet genoeg data aanwezig was. We hebben daarom besloten om puur te kijken naar de voedselprijzen en conflicten in deze regio. Voor de conflicten in deze regio hebben wij een eigen dataset gemaakt, met behulp van gegevens van de bbc (www.bbc.com). Ook hebben wij besloten ons eerst te focussen op de verplichte vragen, voordat wij ons echt gaan focussen op onze eigen vragen. Dit om het verplichte gedeelte in ieder geval af te hebben, om te voorkomen dat wij er later achter komen dat deze vragen meer tijd kosten dan verwacht.

Deel 2:

Na veel onduidelijkheid over wat wel en niet zou mogen op de demo, hebben wij besloten geen gebruik te maken van moeilijke interactieve grafieken. Hoewel ons dit wel veel mooier en leuker leek, vonden wij het te risicovol om heel veel tijd te steken in een moeilijke techniek, terwijl het nog niet zeker was of wij deze ook daadwerkelijk zouden kunnen laten zien. Toen we dit eenmaal hadden besloten, zijn we ons gaan focussen op hoe we onze data dan wel wilden representeren. We hebben toen besloten eerst individueel aan de slag te gaan met specifieke grafieken, om te ontdekken wat wel en niet mooi zou zijn.

Week 3

Maandag:

Om meer tempo te maken is er besloten om de vragen te verdelen zodat iedereen een duidelijk doel had om naar toe te werken. Verder hebben we besproken wat voor een grafieken de data het beste zouden representeren. Voor onze vragen en data was het vooral de vraag of wij voor lijngrafieken of scatterplots zouden kiezen. Omdat de missende data de interpretatie van de lijngrafieken sterk beïnvloedde, hebben wij besloten om voornamelijk met scatterplots te werken. Echter zijn lijngrafieken niet helemaal verbannen, als iemand het gevoel had dat deze nodig waren, konden deze alsnog toegevoegd worden.

Dinsdag:

Discussie over website gehad. Meningeën verschilden, maar niemand had al echt een goed beeld van wat het beste zou zijn, dus besloten om hier nog langer over na te denken.

Woensdag:

Onderling code uitgewisseld om meer richting in stijl grafieken te gaan. Besloten het simpel te houden, om het overzichtelijk te houden en bugs zoveel mogelijk te voorkomen.

Donderdag:

Lang gesprek gehad over technisch rapport/logboek. Besloten dat iedereen eigen analyses maakt maar dat er samen wordt gekeken naar zaken als inleiding/methode. Logboek voor persoonlijke progressie aangemaakt, om Process Book op orde te krijgen. De genoemde progressie in het drive bestand vonden wij niet genoeg om een mooi process book te schrijven.

Vrijdag:

Besloten wat wij wel en wat wij niet op website willen. Iedereen was het eens dat er een interactieve wereldkaart zou komen. Onenigheid over of er informatie per conflict zou moeten komen, besloten om het niet te doen omdat de invloed van de conflicten nog niet duidelijk genoeg te zien was. We dachten dat een speciale pagina met informatie per conflict de indruk zou wekken dat deze conflicten d oorzaak zouden zijn van de bijgeleverde informatie over voedselprijzen, terwijl dit natuurlijk niet altijd de oorzaak hoeft te zijn.

Weekend:

In het laatste weekend hebben wij als groep niet meer samen afgesproken. Via whatsapp hebben wij elkaar op de hoogte gehouden van elkaars process, maar we hebben geen grote beslissingen meer genomen.

Week 4

Maandag:

We hebben besloten om gebruik te maken van een toggle-functie. Om conflicten in een grafiek aan te wijzen leek dit ons onvermijdelijk. Door problemen met de hovertool, en om consistent te blijven met elkaars grafieken, leek ons dit de beste optie.

Dinsdag:

Besloten hoe we de grafieken uiteindelijk echt willen inleveren. Geen discussie over gehad, gewoon besproken dat hoe we het nu hadden, het ook echt gaat worden.

Woensdag:

Besloten wie welk deel van het technisch rapport en het process book zou afmaken. Laatste loodjes verder afgemaakt maar geen grote beslissingen genomen.

Donderdag:

Laatste lelijkheden/onduidelijkheden van website aangepast/verwijderd. Voornamelijk beziggehouden met afronden van het project en geen grote beslissingen meer gemaakt als groep.

Vrijdag:

Demo

Datum	4 Juni - 6 Juni	7 Juni - 10 Juni	11 Juni - 13 Juni	14 Juni - 17 Juni
Persoon/weekdeel	<i>Week 1, deel 1</i>	<i>Week 1, deel 2</i>	<i>Week 2, deel 1</i>	<i>Week 2, deel 2</i>
Tijmen	Dataset verkennen. Git, Python en dergelijke opnieuw onder de knie krijgen en plannen maken voor het onderzoek.	Dataset gepreprocessed zodat alleen relevante landen overbleven in de dataset	Bokeh installeren en beginnen met grafieken maken. Onderzoeksvragen verder uitwerken	Github pages opzetten en eerste files aanmaken.
Thijs	De dataset ontdekken. Begin code schrijven voor preprocessen dataset. Taakverdeling + planning week 1 maken. Git repository opzetten.	Preprocessen dataset voltooien (nutteloze informatie verwijderen, missende data markeren). Werken aan Deliverables week 1.	Leren omgaan met Bokeh om mooie grafiekjes te maken. Externe dataset preprocessen en compatible maken voor eigen dataset.	EDA toepassen op dataset + analyse betekenis visualisatie en hypothesen opstellen. Begin maken aan beantwoorden van de vragen.
Bart	Dataset gekozen. Dataset vervolgens bekeken en externe dataset aangemaakt	Conflictendatabase aangemaakt en conflicten opgezocht. Planning voor komende weken gemaakt.	Introductie met bokeh. Oefenen met plotten met bokeh. Git werkend gekregen. Verder gegaan met conflicten opzoeken	Er voor gekozen om BBC tijdlijn aan te nemen als conflictendatabase. Begin gemaakt met beantwoorden vraag 2
Enno	Kennis gemaakt met dataset. Gezocht naar externe datasets. Taakverdeling en planning gemaakt.	Overzicht gemaakt van geleverd werk/ plan van aanpak voor nieuwe week. Stof doorgenomen en hiermee geoefend met de dataset.	Overzicht en planning gemaakt. Begin gemaakt aan het beantwoorden van de eerste vraag.	Stof doorgenomen en toegepast in praktijk om vraag te beantwoorden. Hypothesen opgesteld en eerste conclusies getrokken.

Datum	18 juni	19 juni	20 juni	21 juni
Persoon/weekdeel	<i>Week 3, maandag</i>	<i>Week 3, dinsdag</i>	<i>Week 3, woensdag</i>	<i>Week 3, donderdag</i>
Tijmen	Fatsoenlijke div structuur gemaakt voor alle HTML code	CSS stijl verbeterd	Side navigation menu aangemaakt met bijpassende icons.	Wereldkaart toegevoegd via javascript
Thijs	EDA deel voor technisch rapport schrijven. Fout in code preprocessen data gefixt (nam units van producten niet goed mee)	Begin kennismaken met sciplearn om regressie en clustering te kunnen toepassen op de data	regressie en clustering verwerken in de code om simpele graphische weergeven te kunnen maken. Begin code schrijven voor de uiteindelijk analyse Vraag3 + extra analyse over conflict gebieden.	Optimaliseren code + functionaliteit voor alle land/product combinaties toevoegen.
Bart	Begin met programmeren van vraag 2. Uitgezocht welke landen en producten gebruikt zullen worden	Vraag 2 qua programmeren werkend gekregen	Legenda werkend gekregen. Verder onderzocht welke landen te gebruiken voor vraag 2.	Scatterplot en clustering geprobeerd te maken
Enno	Logboek aangemaakt. TODO lijst gemaakt/ verdeeld.	Vraag1 verder beantwoord. Overzicht van bruikbare data gemaakt.	Nieuwe stof bekeken/ geoefend.	Plan van aanpak voor technisch rapport/ vragen beantwoorden. Logboek gemaakt.

Datum	22 juni	23 juni	24 juni	25 juni	26 juni	27 juni	28 juni	29 juni
Persoon/weekdeel	<i>Week 3, vrijdag</i>	<i>Week 3, zaterdag</i>	<i>Week 3, zondag</i>	<i>Week 4, maandag</i>	<i>Week 4, dinsdag</i>	<i>Week 4, woensdag</i>	<i>Week 4, donderdag</i>	<i>Week 4, vrijdag</i>
Tijmen	Graph gemaakt met stijl die overeenkomt met de stijl van de website	Alle scripts geoptimaliseerd en HTML code juist geformat	Lijst van conflicted verwerkt van excel naar HTML table	conflicentabel onderverdeeld in landen.	Lijst met alle landen voor alle conflicten vervangen met vlaggen	Wereldkaart verwijderd uit sidebar. Neergezet op pagina graphs en meer toegevoegd	Alle grafieken op de website gezet. Puntjes op de l.	DEMO
Thijs	Opzet inleiding technisch rapport opstellen. Code voor analyse vraag 3 afmaken (volledige functionaliteit + machine learning technieken + interactieve visualisaties)	Weekend weg	Weekend weg	Genereren grafieken voor conflict landen (eigen analyse en vraag 3) voor correlatie tussen data en externe data. Code uitbreiden met non-graphical analysis.	Non graphical analysis toevoegen aan de grafieken. Code afmaken voor correlatie relateren aan conflicten binnen de geanalyseerde landen	Inleiding technisch rapport afmaken + methode technisch rapport afmaken	afronden verslag: resultaten + laatste deel discussie	DEMO
Bart	Correlaties tussen landen opgezocht en scatter plots	Span met conflicten annotaties proberen te maken	Gekozen voor uiteindelijk plots.	Analyses voor vraag 2. Hypothesen en conclusie opgeschreven	Span en buttons werkend gekregen. Layout grafieken afgemaakt	Laatste aanpassingen aan grafieken. Verslag	Verslag afronden. Project inleveren	DEMO
Enno	Begin van technisch rapport gemaakt.	Hard-coding zoveel mogelijk uit code gehaald. Plan van aanpak voor laatste week opgesteld.	Alle plots gemaakt/ gezocht die we willen gebruiken voor analyse	Fouten eruit gehaald en begin analyses. Technisch rapport update.	Analyses en klaar-maken grafieken voor website.	Analyses afronden en kleine aanpassingen aan grafieken. Technisch rapport	Afronden verslag/ afronden process book: afronden project.	DEMO

Screenshots van logboek