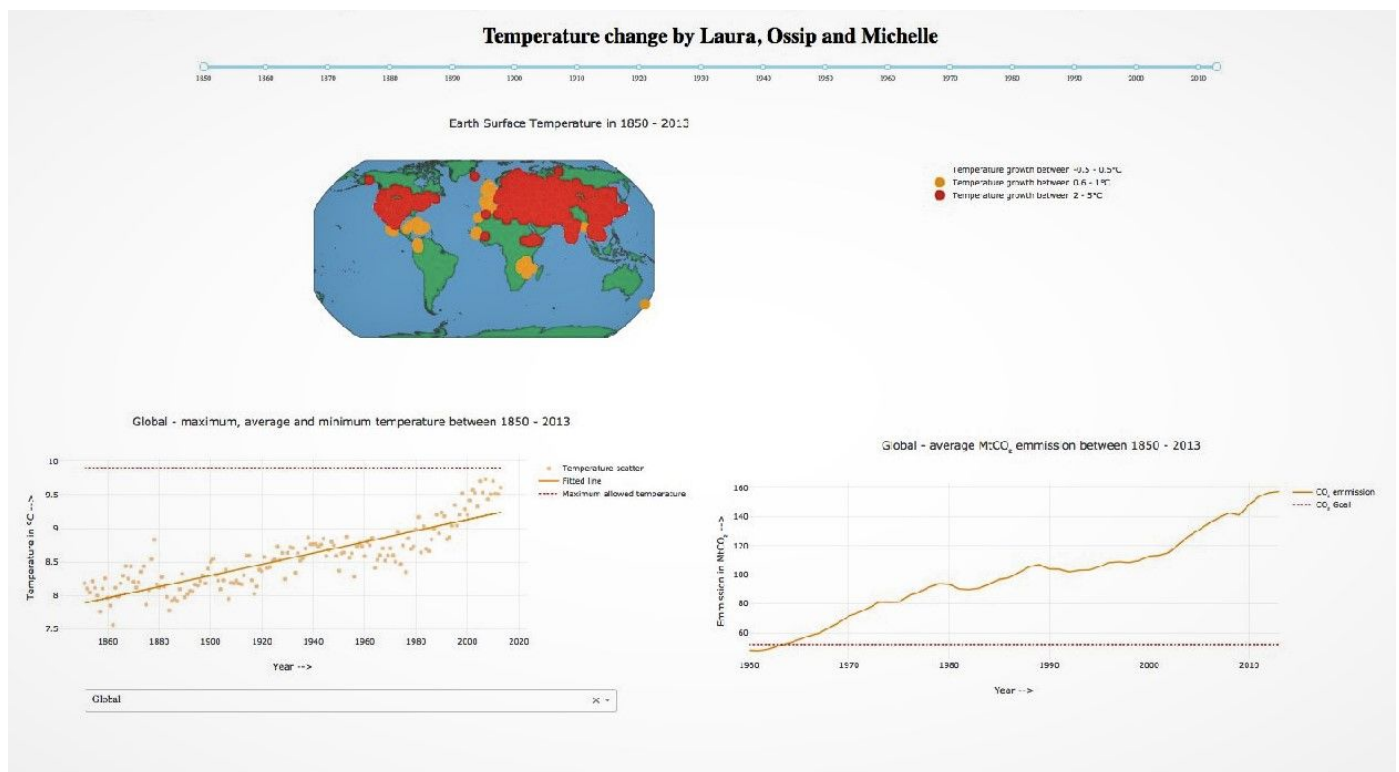


Informatievisualisatie

Projectrapport

Laura Hilhorst, 11048999
Michelle Frankhuizen, 10365354
Ossip Kupperman, 11013583

Geschreven op 28/06/2018



Introductie

Sinds de Industriële Revolutie omstreeks 1750 is onze maatschappij veranderd naar een consumptiemaatschappij waar (in de Westerse wereld) schaarste steeds minder werd. We hebben de productiekracht gekregen om in rap tempo enorm veel spullen te produceren, verkopen en vervoeren. Dit heeft ervoor gezorgd dat ons leven een stuk comfortabeler is geworden: in plaats van ons eigen vee te houden, onze eigen garen te spinnen, truien te breien en manden te weven kunnen we tegenwoordig bijna alles online bestellen en laten bezorgen tot aan de voordeur. Ook kunnen we in luttele uren de hele wereld over reizen en hoeven we ons nooit zorgen te maken om kou in huis, het vergaan van ons eten of het bereiden daarvan.

Om al deze technieken mogelijk te maken, vragen we echter veel van de aarde: fabrieken, auto's en elektriciteit worden namelijk aangedreven door fossiele brandstoffen. Niet zo lang geleden zijn we er echter achtergekomen dat deze brandstoffen, naast dat de voorraad eindig is, een groot effect hebben op ons klimaat. De uitstoot die ontstaat bij de verbranding van fossiele brandstoffen zorgt ervoor dat er onder andere meer koolstofdioxide en methaan in de ozonlaag terecht komt, waardoor de uitstraling van warmte door de aarde beter wordt vastgehouden. Hierdoor warmt de aarde in zijn geheel op, wat drastische gevolgen heeft voor het klimaat. Als bijvoorbeeld de poolkappen smelten, dan stijgt de zeespiegel en ontstaan natuurverschijnselen op plaatsen waar ze eerst niet voorkwamen. Winters worden milder, waardoor insectenplagen vaker voorkomen, oogsten mislukken door extreme of uitblijvende regenseizoenen, overstromingen worden gebruikelijker, enzovoorts. Deze gevolgen kunnen voor landen met minder goede infrastructuur en gezondheidszorg leiden tot enorme humanitaire crises.

Klimaatverandering is een omstreden onderwerp en je zult er ook geen eentonig antwoord over krijgen van verschillende partijen. Anno 2018 is er voldoende wetenschappelijk bewijs om te bevestigen dat we in een stijging zitten die groter is dan de natuurlijke fluctuaties van de aarde maar omdat een heel groot deel van onze wereldeconomie gebaseerd is op het verbranden van fossiele brandstoffen, krijgen nieuwe, schonere ondernemingen veelal de wind van voren. Daarnaast voeren bedrijven die baat hebben bij het gebruik van fossiele brandstoffen hun eigen onderzoeken uit om global warming tegen te spreken.

Toch zijn de meeste overheden het met elkaar eens dat er iets moet veranderen: In 2015 is in Parijs het Klimaatakkoord door 195 landen getekend. Deze landen hebben de uitvoering van dit akkoord op nationaal niveau laten uitvoeren. Er is geen vastgestelde einddatum voor de uitvoering van de volgende speerpunten van het akkoord:

- De globale temperatuurstijging moet onder de 2 graden Celsius blijven ten opzichte van pre-industriële temperaturen, maar onder de 1,5 graden is gewenst.
- In 2030 moeten we wereldwijd 49% minder broeikasgassen uitstoten dan in 1990.

Deze waarden zijn dan ook de toetsingspunten van onze visualisaties. Het doel van de visualisaties is om te laten zien hoe de wereld er aan toe is: we zien de uitstoot en de opwarming van de aarde en hoe drastisch deze veranderd zijn vanaf de industriële revolutie. Ook zien we hoe ver we van onze doelen zitten. Het heeft ons meer dan 250 jaar gekost om deze verandering zo heftig te laten worden, maar we hebben slechts decennia om het weer terug te draaien. De hoop is om met deze visualisaties te laten zien hoe hard we aan het werk moeten om deze doelen te behalen, om een wereldwijde crisis te vermijden.

Voorgesteld Dashboard

De eerste dataset bestaat uit data van 1750 tot 2013, gemeten over bijna de hele wereld. Deze bestaat uit temperaturen, gemeten per jaar per stad en per land, voor zo ver metingen gedaan zijn. In de 18e eeuw was namelijk niet overal ter wereld de apparatuur beschikbaar om deze metingen te doen, of zijn de metingen niet geregistreerd.

De csv bestanden uit deze dataset die voor onze visualisatie gebruikt worden, zijn de bestanden 'GlobalLandTemperaturesByCity.csv' en 'GlobalLandTemperaturesByCountry.csv'. Deze bestanden geven de gemiddelde temperatuur weer per stad en per land met data vanaf 1750 tot 2013. De grootte van 'GlobalLandTemperaturesByCity.csv' is ongeveer acht miljoen rijen met zeven kolommen. Het inladen van deze dataset duurt daarom vrij lang. De gegevens in de verschillende kolommen zijn onzekerheid over de gegevens, de gemiddelde temperatuur, met daarbij de lengte- en breedtegraad per locatie.

Om onze visie op dit onderwerp uit te drukken, is het noodzakelijk om verschillende aspecten van klimaatverandering te betrekken. Om deze reden is een andere dataset betrokken bij de visualisatie, die de koolstofdioxide emissie weergeeft per land. Deze dataset verschilt aanzienlijk van de originele dataset met waarden van temperatuur. De dataset voor temperatuur bevat kolommen met de volgende gegevens: datum (dd-mm-yyyy), lengtegraad, breedtegraad en onzekerheid. De koolstofdioxide emissie dataset loopt van 1960 tot 2016. Deze gegevens zijn globaal en hebben de functie om de uitstoot van broeikasgassen binnen een bepaald tijdsbestek weer te geven. Het is een combinatie van uitstoot door cement, uitlaatgassen, dieren en andere factoren op globaal gebied.

De data wordt bij de twee datasets op verschillende wijzen ingeladen en gesorteerd. De dataset voor temperaturen wordt eerst ontdaan van lege en foutieve waarden, omdat deze geen toevoeging bieden aan de gekozen visualisaties. Vervolgens wordt voor elke stad en land de gemiddelde temperatuur berekend door het gemiddelde van de 12 maanden in een jaar te nemen. Dit heeft meerdere redenen: de data heeft dan minder uitschieters heeft vanwege winter- en zomertemperaturen en de visualisatie heeft uiteindelijk minder computerkracht nodig bij het plotten van de visualisatie.

De derde dataset bevat informatie over de koolstofdioxide-uitstoot per land en per jaar. Om deze data te gebruiken in de globe moet wederom de lege en foutieve waarden verwijderd worden. Tot slot moet de koolstofdioxide-uitstoot over de hele wereld bepaald worden in plaats van per land.

Onze visualisatie bestaat uit drie componenten. Het eerste component is een visualisatie met behulp van een wereldkaart. Door het gebruik van de globe krijgt de gebruiker een beter beeld van de kolommen die de locatie betreffen binnen de dataset. Aan de hand van een globe worden verschillen in gemiddelde temperatuur per stad per jaar weergegeven. Dit wordt gedaan aan de hand van een slider, die de gebruiker door de verschillende jaren binnen de dataset laat navigeren. Deze slider gebruikt twee waarden, een beginwaarde en een eindwaarde, waar het verschil in temperatuur tussen wordt gemeten. Door middel van een gradiënt, die van blauw (een negatief verschil in temperatuur) naar rood (een groot positief verschil in temperatuur), worden de verschillen weergegeven op de wereldkaart. Wanneer er bij een meting geen verschil aangetroffen is, wordt een witte omlijning gebruikt ter verduidelijking. Deze gradiënt wordt gekoppeld aan 'bubbels' op de kaart.

Het tweede component van de visualisatie is een grafiek van de gemiddelde, minimale en maximale temperaturen per land. De standaard weergave betreft een globale view. Deze view neemt metingen van alle landen binnen een bepaald tijdsbestek en betreft altijd jaren. Dit tijdsbestek kan gekozen worden door de gebruiker van het dashboard met behulp van een slider. Zo kan de gebruiker de globale temperatuursveranderingen selecteren en vergelijken met andere tijdsperiodes. Dit geeft meerdere dimensies weer van globale veranderingen op het gebied van gemiddelde temperatuur. Naast de globale weergave kan de gebruiker aparte landen selecteren met behulp van een dropdown. Hierdoor kan de gebruiker naar eigen keuze landen selecteren die hem of haar interesseren. Daarbij worden de verschillen in extreme temperaturen ook duidelijk gemaakt aan de gebruiker, aangezien deze toonaangevend zijn voor het klimaat.

In deze grafiek wordt ook de maximaal toegestane temperatuur (2 graden Celsius boven pre-industriële temperatuur), zoals afgesproken in het klimaatakkoord, weergegeven. Hierbij is het wel belangrijk te melden dat er geen afgebakende definitie is voor “pre-industriële” temperaturen. Omdat er pas vanaf 1850 wereldwijde temperatuur metingen werden opgeslagen, wordt de periode 1850-1900 doorgaans als referentiepunt genomen. Zo ook in onze visualisatie.

Op deze manier krijgt de gebruiker een idee van de stijging in temperatuur en de toegestane limiet voor deze stijging.

Het derde component van de visualisatie is een lijngrafiek van de koolstofdioxide-emissie in metrische tonnen per jaar. De afgesproken waarde van het klimaatakkoord (49% van de uitstoot ten opzichte van 1990) ook verwerkt in deze visualisatie als een lijn. In tegenstelling tot de twee andere componenten geeft de koolstofdioxide emissie een globaal beeld. Hiermee wordt een contrast gegeven tussen lokale temperatuurstijgingen, per stad of per land, tegenover de globale stijging van de uitstoot van koolstofdioxide.

Evaluatie van het dashboard

Het doel van informatievevisualisatie is het ontdekken van de structuur van een dataset. Deze structuur is aan de ontvanger van de visualisatie nog niet bekend. Volgens dit doel is de visualisatie pas succesvol, wanneer deze structuur onthuld wordt door het interacteren en zien van de visualisatie. Tijdens de ontwikkeling van het dashboard zijn wij door het proces gegaan dat ons is geleerd tijdens dit vak. Eerst wordt een dataset verkend, daarna beredeneerd wat er het beste mee gedaan kan worden en vervolgens gepresenteerd op een begrijpelijke wijze.

In eerste instantie waren we overweldigd door de omvang van de datasets, maar na enige verkenning en begripsontwikkeling van de gegevens binnen de dataset, ontstond het besef dat er verschillende manieren waren om deze gegevens te visualiseren. Een beeld geeft immers meer weer dan duizend woorden. Met behulp van de kolommen ‘longitude’ en ‘latitude’, ook wel lengte- en breedtegraad genoemd, was het mogelijk om locaties weer te geven. Aan de hand van de kolom met dateringen van de metingen van temperatuur kon een tijdsperiode worden weergegeven. Dit kon niet in een traditionele grafiek gedaan worden, aangezien gegevens voor locatie dan irrelevant werden.

In de lijn van het verhaal dat door onze visualisatie verteld wordt, zouden meer aspecten van klimaatverandering aan bod moeten komen dan temperatuurstijgingen. De uitstoot van koolstofdioxide draagt ten zeerste bij aan klimaatverandering en kan worden gekoppeld aan lokale

en globale stijgingen in temperatuur. Om deze reden is een externe dataset toegevoegd met de globale emissie van koolstofdioxide door de jaren heen. Deze dataset kwam echter niet geheel overeen met de structuur van de originele dataset; er zijn pas vanaf 1960 regelmatige en betrouwbare metingen naar koolstofdioxide-uitstoot gedaan.

Als presentatievorm is voor een lijngrafiek gekozen, aangezien deze grafiek de ernst van de situatie duidelijk weergeeft en de indruk wekt van een continu proces, wat klimaatverandering absoluut is. Een interessant vervolgonderzoek zou zijn om op basis van deze waarden voorspellingen te doen voor de toekomst en de invloed die de afspraken van individuele landen hebben op de uitkomst.

De presentatievorm van het temperatuurverschil per stad is met behulp van de lengte- en breedtegraad gevisualiseerd als een wereldkaart. Hiermee wordt een beeld geschetst van de wereld en in hoeverre deze verandert, of veranderd is, met de jaren op het gebied van temperatuur. Er wordt getoond dat serieuze stijgingen in temperatuur aanwezig zijn over bijna de gehele wereld. Als men zich aan het klimaatakkoord wil houden en de aarde wil behouden als veilige leefomgeving, dan zullen drastische maatregelen getroffen moeten worden.

De dataset voor metingen in temperatuur voor grote landen is anders ingedeeld dan die per stad. Deze geeft geen precieze locatie, maar alleen de naam van het land. Hierdoor is gekozen voor een lijngrafiek van temperaturen door de jaren heen. Door het geven van minimale, maximale en gemiddelde temperaturen worden verschillen tussen en stijgingen in extremen duidelijk gemaakt aan de gebruiker. Het door ons vertelde verhaal wordt hierdoor alleen maar versterkt.

Reflectie op het Teamwerk

Vorbereidingsfase

Als team zijn we door alle mogelijke onderwerpen heengelopen voor deze opdracht. Na uitvoerig discussiëren kwamen we uit op twee mogelijkheden. Het verhaal dat wij zouden vertellen over klimaatveranderingen enthousiasmeerde ons meer dan het verhaal over terrorisme. Daarom zijn we dit idee verder gaan uitbouwen. Na het inleveren van onze individuele opdrachten zijn we harder gaan nadenken over de mogelijkheden van onze dataset. Door de tools te gebruiken die ons zijn aangereikt tijdens de individuele opdrachten, kregen we ook een beter beeld over onze mogelijkheden voor een mooie visualisatie. We hebben gespeculeerd over de opmaak, wat de gebruiker zou denken als deze het zag en waarom de gebruiker bepaalde dingen zou doen. Daarna zijn we op onze visie gekomen, die samenhangt met het in 2015 ondertekende klimaatakkoord in Parijs. Vervolgens hebben wij besloten hoe de data gegroepeerd zou worden en welke datasets door ons gebruikt zouden worden. Na alle eigenschappen van de gegevens geëvalueerd te hebben, werden eerder genoemde datasets door ons gekozen voor het dashboard.

Ontwikkelfase

De ontwikkelingen binnen ons project liepen gestaag. Gezamenlijk kwamen we er goed uit en als er enige problemen waren, dan besproken we die met elkaar. Bij het programmeren nam Michelle het voortouw, aangezien zij daar veel affiniteit mee heeft. Zij was als het ware de technisch directeur en hielp de anderen wanneer zij ergens tegenaan liepen. In week drie zijn we daadwerkelijk aan de slag gegaan met het bouwen van het dashboard en het combineren van de gemaakte grafieken. Hierbij liepen we tegen een aantal problemen aan, zoals het updaten van grafieken door middel van user-input. Door dit met elkaar te bespreken en de student assistent om hulp te vragen zijn we

hieruit gekomen. Na het maken van de eerste twee visualisaties ontstond het gevoel dat alleen temperatuurmetingen een te oppervlakkig beeld gaven van klimaatverandering. Daarom zijn wij in week drie op zoek gegaan naar een extra dataset, zodat de derde visualisatie meer inzicht kon geven. De derde dataset is niet geheel vergelijkbaar met de originele datasets omdat de metingen pas vanaf 1960 gedaan zijn, maar wij achtten hem toch nuttig genoeg om hem toe te voegen.

Taakverdeling

Gezamenlijk

Onderwerp kiezen

Onderzoek doen naar maatschappelijke waarde en relevantie

Brainstormen over visualisaties

Laura

README geschreven

Slider & Dropdown in Dash programmeren

Nakijken rapport, grafieken en Dash

Slides presentatie

Ossip

Project Pitch uitgeschreven

CO2 Lijngrafiek programmeren

Rapport geschreven

Nakijken grafieken en Dash

Michelle

Bubble Map programmeren

Lijngrafiek programmeren

Dashboard samenstellen

Nakijken rapport en readme

Bronnen

Klimaatakkoord Parijs, gevonden op internet op 28/6/2018,

https://nl.wikipedia.org/wiki/Akkoord_van_Parijs