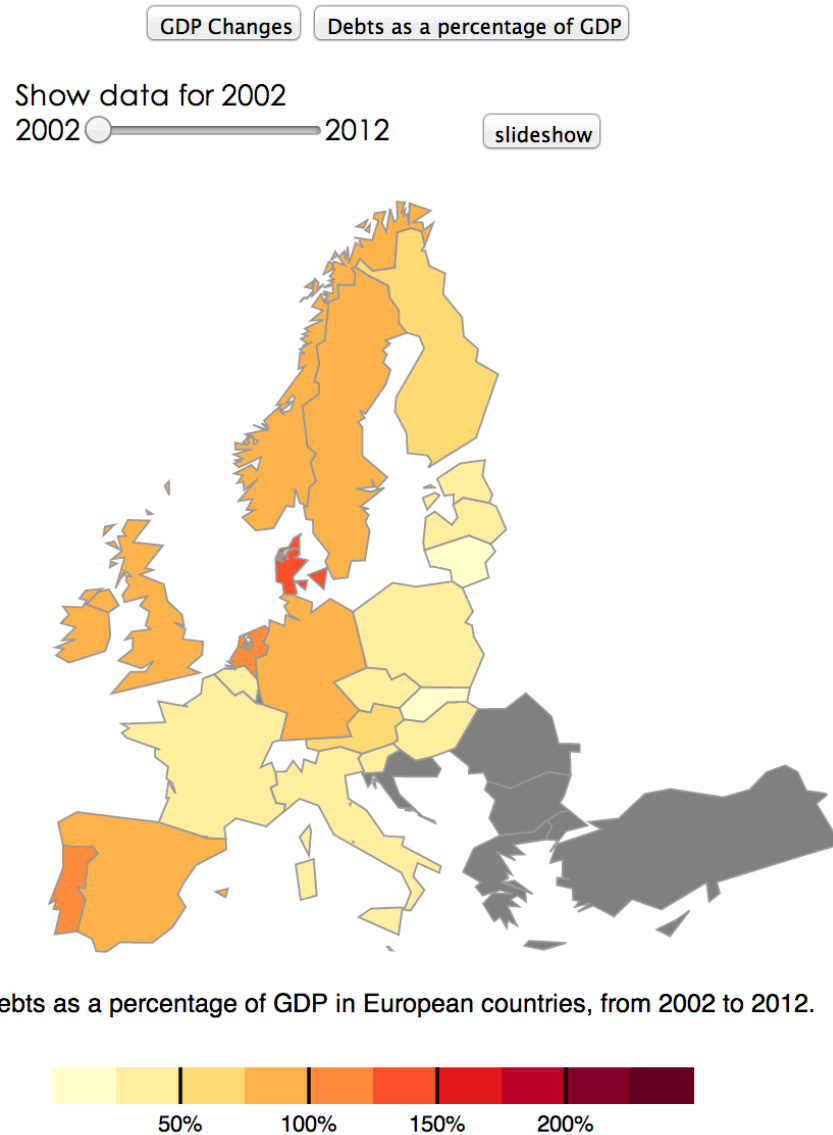


Informatievisualisatie (5072INFO6Y)
D3 Project Eindverslag

Groep 11

Patrick Akil, 10668411
Kelvin Chan, 10634061
Thom de Leu, 10611045
Xavier Marseille, 10564381
Jan Jaap Meijerink, 10560459

26-06-2014



1. Inleiding

Voor het groepswork moesten wij een visualisatie maken over het BNP per hoofd van de bevolking en de verhouding van de schulden ten opzichte van het inkomen. Hierbij moesten wij kijken of er sprake is van een correlatie tussen deze twee gegevens.

Omdat wij het BNP en de schulden ten opzichte van het inkomen in Europa moesten visualiseren vonden wij het handig de kaart van Europa te gebruiken. Wij waren van mening dat dit voor onze opdracht de duidelijkste visualisatie zou opleveren. Het leek ons duidelijker omdat iedereen het beter voor zich ziet. Je kunt eerder verbanden leggen tussen bijvoorbeeld buurlanden omdat je de kleuren met elkaar kan vergelijken. Door bijvoorbeeld alleen namen en percentages te geven zullen mensen minder snel eventuele verbanden kunnen leggen.

Bij de visualisatie hebben wij er voor gekozen om de twee gegevens die wij moeten visualiseren gescheiden te houden. Bij de eerste visualisatie over het BNP hebben wij gekozen voor de kleuren rood, licht bruin, bruin, licht groen en groen. De rode en bruine kleuren geven aan dat het BNP in 2012 lager is geworden ten opzichte van 2001. De groene kleuren geven aan dat het BNP in de jaren 2001 tot 2012 is toegenomen. Hierbij geldt hoe donkerder de kleur des te groter de verandering.

Bij de andere visualisatie hebben wij gekozen voor de kleuren van licht bruin tot aan donker rood. Ook hierbij geldt hoe donkerder de kleur des te groter is het percentage van schuld ten opzichte van het inkomen. Bij deze visualisatie hebben wij daarnaast een slideshow toegevoegd zodat men eenvoudig kan kijken naar het verloop in de afgelopen jaren. Hierdoor kan men snel zien of er een grote verandering heeft plaats gevonden. Daarnaast hebben wij ervoor gezorgd dat als er met de muis over een willekeurig land gaat, dat de percentages van het desbetreffende land worden weergegeven.

Wij hebben er daarnaast voor gezorgd dat men ook eenvoudig kunnen zien welke landen in een bepaalde percentage groep vallen. Als men met de muis over deze balk zweeft, zullen de landen die daarbij horen duidelijk zichtbaar worden in de kaart. Daarnaast staan de namen van deze landen naast de kaart aangegeven voor extra duidelijkheid.

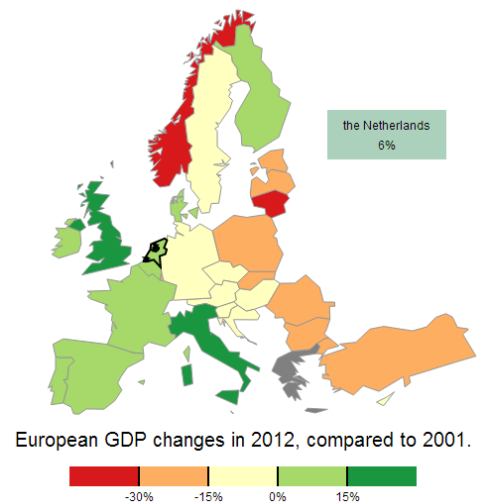
De visualisatie die wij hebben gemaakt is voornamelijk bedoeld voor economen. Het is voor ons van groot belang dat de visualisatie eenvoudig te begrijpen is. Ondanks dat de visualisatie voornamelijk voor economen is bedoeld, vinden wij het ook van groot belang dat anderen het principe van de visualisatie gebruiken. Om die reden hebben wij geprobeerd alles zo simpel mogelijk te houden. Doordat wij niet meer informatie geven dan nodig is wordt de visualisatie niet onoverzichtelijk. Daarnaast is het zo dat mensen met weinig kennis over bijvoorbeeld het bruto nationaal product wel kunnen snappen wat de verandering inhoud.

2. Gerelateerd werk

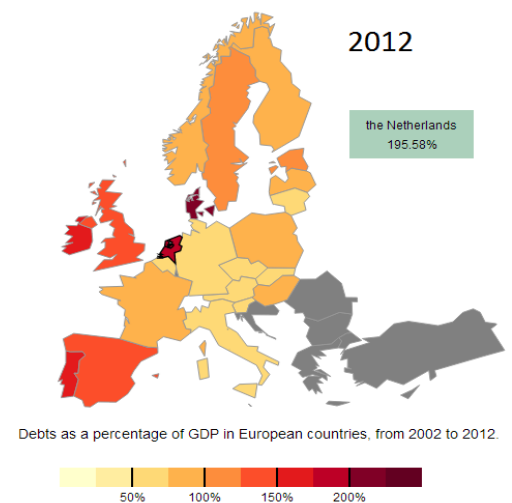
Wij hebben gekozen om de data te visualiseren met een choropletenkaart. Een choropletenkaart wordt als volgt beschreven *“A choropleth map is a thematic map in which areas are shaded or patterned in proportion to the measurement of the statistical variable being displayed on the map, such as population density or per-capita income.”*¹ Kort gezegd: een choropletenkaart is een thematische kaart waarin gebieden in verschillende contrasten worden aangeduid, gebaseerd op de statistische gegevens die erachter liggen.

J. Heer, M. Bostock en V. Ogievetsky hebben in hun artikel ‘A Tour Through the Visualization Zoo’ gezegd: *“Select effective visual encodings to map data values to graphical features such as position, size, shape, and color.”*² Het is voor een informatieve visualisatie belangrijk dat er gelet wordt op de positie, grootte, vorm en kleur van de gemaakte visualisatie. Zeker in het geval van een kaart is het belangrijk om een goed onderscheidt te kunnen maken tussen de verschillende gebieden.

Wij wilden eerst kiezen voor de kleuren rood en groen op al onze kaarten. Deze keuze wilden wij eerst maken omdat rood en groen vrij algemene kleuren zijn in de maatschappij, zo staat groen voor het ‘goede’/‘positieve’ en rood voor het ‘slechte’/‘negatieve’, zoals men bijvoorbeeld bij een stoplicht ook ziet. We zijn uiteindelijk niet volledig door gegaan met deze opstelling. Aangezien we de schuld per huishouden vergeleken met het BNP tussen de jaren 2002 en 2012, moesten de kleuren wel in elkaar over kunnen lopen (zoals u ziet op figuur 2), hierdoor zou de verandering door de jaren heen duidelijker worden. Zoals op figuur 1 te zien is hebben wij hier niet gekozen voor een kleur in hetzelfde contrast, maar toch wel voor groen en rood/oranje, omdat hier binnen één afbeelding de verandering al gevisualiseerd wordt.



Figuur 1

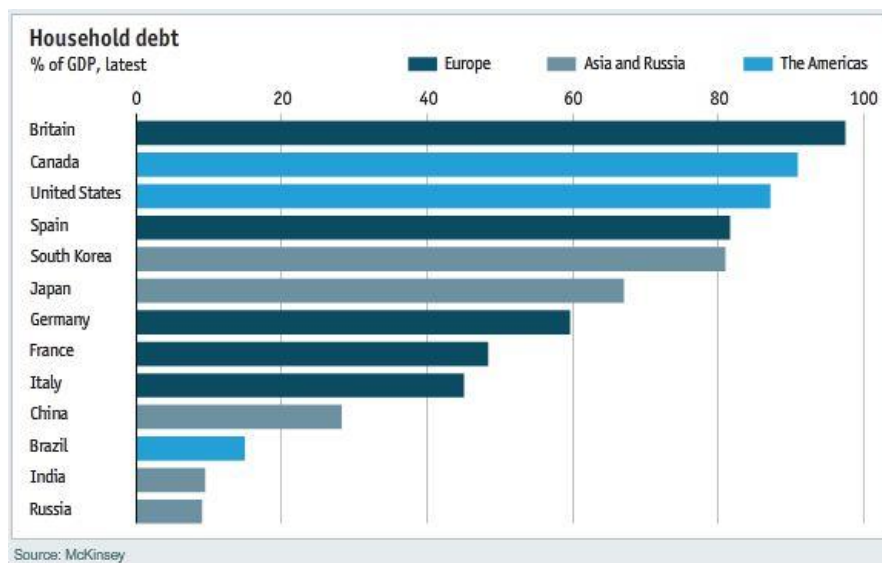


¹ Wikipedia. Choropleth Map. Geraadpleegd op 26-6-2014.
http://en.wikipedia.org/wiki/Choropleth_map

² J. Heer, M. Bostock, V. Ogievetsky. A Tour Through the Visualization Zoo. Communications of the acm, june 2010, vol. 53 no. 6.

Voordat we zijn begonnen met de twee datasets te visualiseren en te combineren hebben we gekeken naar soortgelijke datasets en onderzoeken naar het BNP, het BNP per huishouden, en de schulden per huishouden van een land, een continent of de wereld. We zijn veel verschillende aanpakken tegengekomen, kolomgrafieken, staafgrafieken, lijndiagrammen, et cetera. We hebben drie voorbeelden gekozen waarmee we die van ons kunnen en gaan vergelijken. De, volgens ons, meest toepasselijke voorbeelden. We hebben een staafgrafiek, een lijndiagram en een choropleth map, gecombineerd met een staafgrafiek.

Het eerste voorbeeld gaat over de 'Household debt in percentage of the GDP', oftewel de schuld per huishouden in percentage van het BNP. In dit voorbeeld zijn de grootste dertien landen van de wereld gepakt en zijn de schulden met elkaar vergeleken door middel van een staafgrafiek. Het is in dit geval handig dat ze een staafgrafiek hebben gebruikt, want het is heel overzichtelijk. Er kan zo waargenomen worden welke landen de grootste schuld per huishouden hebben en welke landen de minste. Daarnaast zou in dit geval een choropleth map een lastige optie zijn omdat de landen over de hele wereld verspreid liggen en er dan veel landen die niet van toepassing zijn toch op de kaart zouden staan.



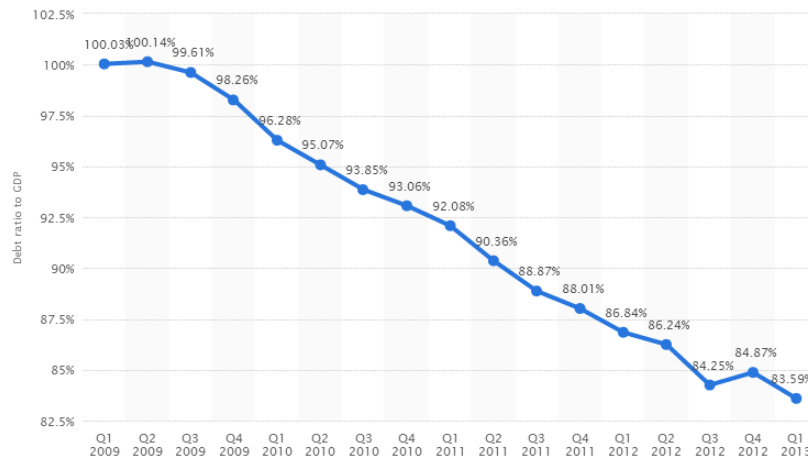
(Figuur 4)³

Als tweede voorbeeld hebben wij de 'Household debt ratio to GDP in the United States between Q1 2009 and Q1 2013'. In dit voorbeeld is gekozen voor een lijndiagram. In dit geval kan dat, aangezien alle data gericht is op één specifiek land, met een lijndiagram kan je direct zien in welk kwartaal van het jaar het beter of slechter ging met het land. Als er zich nou meer landen in de dataset zouden

³ Don't Worry, I'm an economist! Leave economics to the economists! Simple as that. Geraadpleegd op 25-6-2014. <http://im-an-economist.blogspot.nl/2012/03/graphs-of-week-debt.html>

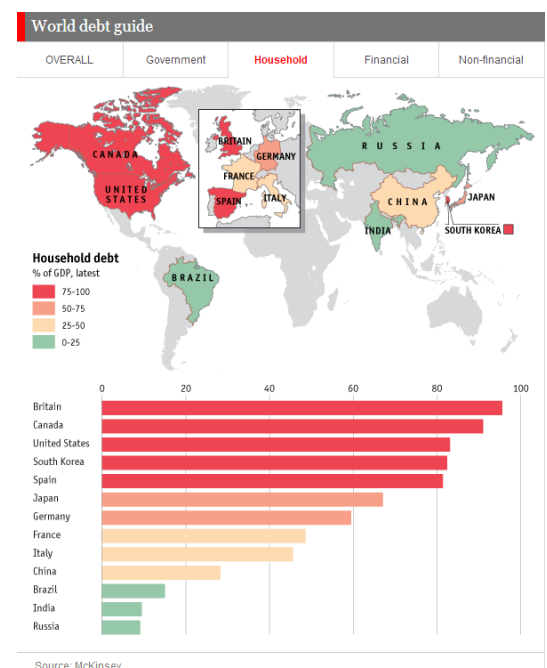
bevinden zou de lijndiagram geen goeie keuze zijn geweest, daarom hebben wij hier ook geen gebruik van gemaakt.

Household debt ratio to GDP in the United States between Q1 2009 and Q1 2013



(Figuur 4)⁴

Het derde vergelijkbare onderzoek waar wij gebruik van hebben gemaakt heeft twee verschillende visualisatie vormen gebruikt, namelijk de choropleth map gecombineerd met een staafgrafiek. Deze kaart/staafgrafiek gaat ook over de schuld per huishouden in percentage van het BNP. Van over de hele wereld zijn weer de grootste dertien landen gekozen om deze met elkaar te vergelijken, net als in voorbeeld een. Nu is er alleen gekozen om naast de staafgrafiek ook nog een choropleth map te plaatsen ter verduidelijking. Net als bij onze map kan er hier over een land gescrold worden om extra informatie te verkrijgen. In het geval van deze dataset is het handig dat ze voor beide hebben gekozen, aangezien de landen niet aan elkaar grenzen en toch wel ver uit elkaar liggen is er een staafgrafiek om te vergelijken, maar de choropleth map is toch het overzichtelijkst.



(Figuur 5)⁵

⁴ The Statistics Portal. Household debt ratio to GDP in the United States between Q1 2009 and Q1 2013. Geraadpleegd op 25-6-2014.
<http://www.statista.com/statistics/248283/household-debt-ratio-to-gdp-in-the-united-states/>

⁵ The Economist. The debtors' merry-go-round. Geraadpleegd op 26-6-2014.
<http://www.economist.com/blogs/graphicdetail/2012/09/global-debt-guide>

3. Voorgestelde visualisatie

Gegevens

De visualisatie is gebaseerd op twee datasets, tec00104.tsv en tec00114.tsv. De inhoud van dataset tec00104.tsv geeft de schuld naar inkomen per huishouden. Deze gegevens zijn opgedeeld per land en een aantal jaren. In dataset tec00114.tsv wordt het bruto nationaal product index (EU28 = 100) per land van een aantal jaren gegeven. Als eerst zijn de data uit de datasets tec00104.tsv en tec00114.tsv geladen. Deze data hebben we een JSON structuur gegeven. Door de JSON structuur is de data overzichtelijker en toegankelijker. Elke object in de JSON is een land met daarbij behorende jaartallen en cijfers van elke dataset. Met de gegevens in de JSON hebben we vervolgens de ratio van de particuliere schulden in percentage van het BNP berekend.

```
console.log(json)
▼ [Object, Object, Object,
  ]
  ▼ 0: Object
    country: "AT"
    ▼ debt: Object
      2002: 77.13
      2003: 76.41
      2004: 79.61
      2005: 84.13
      2006: 85.45
      2007: 85.04
      2008: 86.51
      2009: 85.96
      2010: 89.15
      2011: 89.24
      2012: 86.44
    ▶ __proto__: Object
    ▶ gdp: Object
    ▶ ratio: Object
    ▶ __proto__: Object
  ▶ 1: Object
  ▶ 2: Object
  ▶ 3: Object
```

Visualisatie

Voor onze visualisatie hebben we gebruik gemaakt van een choropletenkaart. We hebben voor onze kleurschalen ColorBrewer geraadpleegd. De ratio's worden op een choropletenkaart van Europa geladen. De landen krijgen volgens hun ratio een bepaalde kleur mee. Deze kleuren staan onder de kaart uitgelegd met een schaalverdeling. Als de muis op een kleur gehouden wordt, dan worden de landen getoond die een ratio hebben in het bereik van de kleur.

Met behulp van de schuifbalk in onze visualisatie kunnen de ratio's per jaar bekeken worden. Ook hebben we een diavoorstelling functie geïmplementeerd. De diavoorstelling functie toont achtereenvolgend de kaarten van een jaartal, beginnend bij 2002 en als laatst 2012. Onze visualisatie heeft twee dimensies.

4. Reflectie op het teamwerk

Vorbereidingsfase

Eerst moesten we weten wat onze opdracht inhield. Toen iedereen wist wat er gedaan moest worden begonnen we te brainstormen over hoe we de opdracht gingen uitwerken. Als iemand een idee had werd dit in de groep gegooid en besproken. Alle groepsleden gaven dan aan of ze het idee goed vonden of juist niet. Hierover werd dan gediscussieerd, dit deden we totdat we een idee hadden dat iedereen goed vond. Hierna gingen we kijken of het idee wel haalbaar zou zijn in de korte tijd die ons toegewezen werd. Na wat onderzoek bleek dat het haalbaar zou zijn als we er gezamenlijk serieus aan gingen werken. De volgende stap was het verdelen van de taken, we wilden ervoor zorgen dat iedereen evenveel aan het project zou doen. Maar om de taken te kunnen verdelen moesten we wel eerst taken hebben. Dus gingen we bedenken wat allemaal precies gedaan moest worden.

We hadden de volgende taken : mid-term verslag, presentatie, coderen zelf voor het project. Dit hadden we toen zo verdeeld:

Patrick Akil – Beginnen coderen

Kelvin Chan – Beginnen coderen

Thom de Leu – Presentatie

Kevin Lok – Presentatie/PowerPoint

Xavier Marseille - mid-term verslag

Jan Jaap Meijerink – Beginnen coderen

Bij het uitvoeren van de verschillende taken zijn er problemen opgetreden door onduidelijkheid over wat precies het idee was. We maakten een planning van wat en wanneer er precies af moest. En iedereen moest zich proberen te houden aan deze planning. Hier zijn ook wat problemen opgetreden omdat niet iedereen zich daadwerkelijk kon houden aan de planning door persoonlijke redenen. Uiteindelijk hebben we de planning hierop aangepast en is alles nog goed gekomen.

Ontwikkelfase

We hadden dus bedacht dat we een kaart wilden gebruiken als visualisatie van onze dataset. Dit was nog een behoorlijke klus om te coderen voor ons. Kelvin en Patrick werkten samen aan de code die alle data filterde en in een JSON zette die makkelijk te gebruiken zou zijn voor de rest van de code. Jan Jaap werkte aan de map zelf en Kevin werkte aan de schuifbalk die de kaart verschillende visualisaties geeft op verschillende tijdstippen. Toen onze aparte codes allemaal werkten, heeft Jan Jaap ervoor gezorgd dat de data uit de JSON van Kelvin en Patrick werd gehaald en goed in de kaart werd gevisualiseerd. Daarna moest ook nog de time slider van Kevin worden geïmplementeerd. Toen dit allemaal werkte begonnen we aan het eindverslag en eindpresentatie. Aan het eindverslag werkten: Kevin, Xavier, Kelvin, Thom. We hadden het eindverslag netjes in stukken verdeeld zodat iedereen een stukje kon doen aan het verslag. Patrick en Jan Jaap werkten in de tussentijd nog aan kleine foutjes in de code. De eind presentatie hebben we gezamenlijk gemaakt. Dit wordt uiteindelijk gepresenteerd door Thom.

Taakverdeling

We wilden de taken zo verdelen dat iedereen evenveel zou doen aan dit project. Dit was de uiteindelijke verdeling van alle taken:

Patrick Akil – Coderen, PowerPoint mid-term Presentatie

Kelvin Chan – Coderen, deel 3 verslag, mid-term verslag, PowerPoint mid-term presentatie

Thom de Leu – Presenteren mid-term presentatie, deel 2 verslag, presenteren eind presentatie

Kevin Lok – PowerPoint mid-term Presentatie, coderen, deel 4 verslag

Xavier Marseille – Website, deel 1 verslag, mid-term verslag

Jan Jaap Meijerink – Coderen

Individuele reflectie op het werk

Patrick Akil: “Op het begin verliep de onderlinge samenwerking niet optimaal. Er ontstond onduidelijkheid over het concept. Deze onduidelijkheid was ontstaan door het gebrek aan communicatie vanuit de groep. Na de mid-term presentatie werden duidelijke afspraken gemaakt. Iedereen had een taak die op een vaste datum afgerond moest zijn. Dit zorgde ervoor dat we niet achter zouden lopen in vergelijking tot andere groepjes. Uiteindelijk ben ik tevreden over het project en denk dat we een goed eindproduct hebben geleverd.”

Kelvin Chan: “Aan de beginfase van ons project was er veel onduidelijkheid. We dachten dat we allemaal hetzelfde concept in ons hoofd hadden, maar toen er resultaat geleverd moest worden, bleek dat toch niet het geval te zijn. Gelukkig kwamen wij hier op tijd achter, waardoor we voor een frisse aanpak gingen. Taken die uitgevoerd moesten worden werden duidelijk verdeeld. Ook moesten de taken helemaal helder zijn voordat er aan begonnen werd. De samenwerking werd daardoor veel beter.”

Thom de Leu: “Onze groep had een vrij langzame start, het was nog een beetje bekijken wat nou precies de opdracht was, wat we precies moesten doen en wat we nou met de gegeven datasets konden maken. Rondom de presentaties kwamen we er snel achter wat we wilden doen en hebben dit in volle gang gezet. De rollen waren in het begin ook nog niet duidelijk afgesproken, maar gaande weg zagen we vanzelf waar onze sterke en zwakke punten lagen en waar we die op konden toepassen. De samenwerking verliep prima en iedereen hield zich gewoon aan zijn taken, ondanks dat we niet altijd een even duidelijke planning hadden.”

Kevin Lok: “Ik vond de samenwerking op het begin wat stroef lopen, er was onduidelijkheid over het project. Iedereen had eigenlijk een ander idee over wat het plan nou eigenlijk was. Dit was het geval voor de mid-term presentatie. Na die presentatie hebben we ons herpakt, we hadden toen duidelijke afspraken gemaakt met elkaar. Hierdoor is het uiteindelijk nog goed gekomen en zijn er geen problemen bij gekomen. Iedereen deed goed zijn best dus van meelifers hebben we geen problemen gehad.”

Xavier Marseille: "Aan het begin van het project was het moeilijk een begin te maken. Het was lastig om voor ons te zien wat wij precies moesten doen en hoe wij het wilden doen. Toen wij eenmaal een idee hadden ging het een stuk beter lopen. Uiteindelijk ging het project redelijk vlekkeloos. Ondanks dat we niet echt een hele strakke planning hadden is er wel onderling duidelijk gepraat over wat er allemaal moest gebeuren. Onderling hebben wij geen problemen gehad, de samenwerking verliep goed. "

Jan Jaap Meijerink: "De samenwerking kwam bij ons erg natuurlijk tot stand, ik deed het meeste van het programmeren, daarom heb ik minder tijd besteed aan bijvoorbeeld de verslagen en de presentaties. Daar hielden de anderen zich dan weer mee bezig. Volgens mij hebben we uiteindelijk allemaal een goede bijdrage geleverd aan het eindresultaat."