

Data processing

Design 4

Groep B3:
Jaap Nieuwenhuizen
Daniël Hansen
Arnoud Smit
Marije Dekker

Problem 1: Design critique group reflection

Take your analysis and ideas from the design critique of last week and discuss them with your group (**Confluence by Harshawardhan Nene and Kedar Vaidya**).

Can you find a consensus?

Op het eerste gezicht is de visualisatie best verwarrend, maar nadat de tekst eronder (meerdere keren) is gelezen snapten we allemaal wat er weergegeven werd. Op het eerste gezicht is het namelijk een hele onduidelijke visualisatie. De opties die eronder staan zijn daarentegen wel heel praktisch. Daarnaast is het contrast tussen roze en paars te laag en wat de witte zandloper nou precies is, is onduidelijk.

Prepare one document in which you in detail critique the visualisation as a group (max 2 pages).

Confluence is een interactieve visualisatie voor Hollywood films. Het geeft een interactieve weergave van een aantal algemene gegevens over alle Hollywood films, zoals de acteurs, kosten, opbrengsten, jaartal, tijdsduur en uiteraard een algemene beschrijving van de inhoud van de respectievelijke film.

De hoofdtak van Confluence is echter om een weergave te geven van de beoordelingen of meningen van 0 tot 100 over Hollywood bioscoopfilms uit 2007 tot en met 2011 van zowel (professionele) critici als het publiek. Dezelfde film wordt in de visualisatie als icoon dus tweemaal getoond: een paarse cirkel betreft de gemiddelde beoordeling van critici, en de roze cirkel van het publiek. Elke film wordt dus weergegeven als een cirkelicoon: beoordelingen van gebruikers/publiek worden weergegeven in roze en beoordelingen van critici in paars. Dit zijn relatief geschikte coderingen voor de datapunten, omdat er relatief veel Hollywood films zijn. De gegevens zijn verzameld via Metacritic. Gebruikers van deze site behoren tot het 'publiek', terwijl beoordelingen van critici geschreven en gepubliceerde recensies betreffen. De gedetailleerdere data over budget, winst, en dergelijke, worden echter niet gegeven.

De achtergrond van de visualisatie wordt gekozen op basis van de laatst geklikte film. Er kan expliciet gezocht worden naar een opgegeven film, er kan gescrold worden over alle films, maar er kan ook handig gebruik gemaakt worden van een aantal functies die het zoeken naar bepaalde films gemakkelijker maakt. De films kunnen gegroepeerd worden op basis van de gemiddelde beoordelingen, en zodoende kan dus geselecteerd worden om enkel films van een bepaalde beoordelingscategorie te tonen (zeer zwak, zwak, gemiddeld, sterk, zeer sterk). Ook kan gekozen worden om films met een bepaalde consensus tussen het publiek en critici te tonen. Zodoende kan gezocht worden naar films waarbij de meningen van critici en publiek sterk verschilt of juist in grote mate in overeenstemming is. Daarnaast kunnen de beoordelingen van ofwel critici ofwel het publiek verborgen worden in de visualisatie, zodat de beoordelingen van enkel de andere groep getoond blijven. Tot slot is

er nog de mogelijkheid om de iconen van de films in grootte te schalen op basis van eigenschappen als budget, opbrengst, inkomsten (wereldwijd, binnenlands, buitenlands) en openingsgegevens. Deze laatste gegevens zijn eveneens de 'gedetailleerde gegevens' die weergegeven worden over een film als geklikt wordt op het icoon van de respectievelijke film.

Aan de linkerkant van de visualisatie vonden de critici de film beter dan het publiek, terwijl dit aan de rechterkant andersom is. Hoe meer naar het midden hoe meer deze twee partijen het met elkaar eens zijn. Het publiek wordt weergegeven met een roze cirkel en de critici met een paarse. De visualisatie kan gefilterd worden op het soort publiek (critici, publiek of beide) en / of op de beoordeling van de film, daarnaast kan er ook gefilterd worden op in hoeverre de critici en het publiek het met elkaar eens zijn. Zelfs het soort film kan gefilterd worden op een aantal categorieën. Wanneer er over een bolletje/cirkel (zowel roze als paars) wordt gegaan met de muis dan verschijnt linksonder een blokje met gedetailleerdere informatie over de film: de naam, verschijningsdatum, het budget, de opbrengst en de exacte beoordelingen van de critici en het publiek. Hiermee kan er gekeken worden naar het verschil tussen critici en het publiek van de films. Op deze manier kan er ook gekeken worden of een film goed of slecht was.

Wanneer er wordt gekeken naar Tufte's principes voor grafische integriteit kunnen de volgende dingen worden geconcludeerd:

- De schalen zijn correct gelabeld, zowel horizontaal als verticaal.
- De lie-factor is niet hoog, er is namelijk geen effect in grootte in de visualisatie.
- De visualisatie laat vooral data variatie zien.

Over Tufte's visualisatie design principes is het volgende te zeggen:

De assen zijn geheel niet gelabeld. Dit maakt dat de visualisatie op het eerste oog niet gelijk helder is. In de tekst linksonder staat echter wel aangegeven wat alles in de visualisatie inhoudt. In de visualisatie is sprake van enkel data variation. Het design van de visualisatie blijft constant: alleen de data verandert. Er is dus sprake van een eerlijke presentatie van de data.

Er is geen sprake van chart junk: alle weergegeven data en informatie heeft een doel. Wanneer een film geklikt wordt, wordt deze gebruikt als achtergrond. Dit is een leuk effect, maar kan enigszins afleidend zijn. Er is een hoge data-ink ratio: voor de data zijn felle kleuren gebruikt (= veel inkt), terwijl de achtergrond en andere elementen aanzienlijk minder fel gekleurd zijn. Er wordt ook handig gebruik gemaakt van layer information: als geklikt wordt op een film, dan worden uitgebreidere gegevens over die respectievelijke film weergegeven in een gelaagd menu. De data densiteit is hoog in deze visualisatie. Volgens Tufte's principes is dit goed, zelf vinden we deze iets te hoog.

De verschillende cirkels zijn bijna niet te onderscheiden en het is lastig iets te herkennen. Dit zou voorkomen kunnen worden door de cirkel uit het scherm te laten springen wanneer er met de muis overheen wordt gegaan. Daarnaast zou het voorkomen kunnen worden door de verschillende cirkels kleiner te maken. Deze zijn namelijk onnodig groot.

Door gebruik te maken van 4 hoofdkleuren is het contrast relatief duidelijk, behalve tussen de visual encodings. Deze zijn namelijk paars en roze. Er kan optioneel gebruik gemaakt worden van grootte als middel om op eigenschappen te onderscheiden. Er wordt geen gebruik gemaakt van vorm, tint, kleur of proximity/nabijheid. De films op de x-as die dicht bij elkaar liggen houden geen verband met elkaar. Wel wordt alignment effectief gebruikt, omdat de beoordeling van het publiek van een bepaalde film op dezelfde plek op de x-as

bevindt als de beoordeling van critici van diezelfde film. Ook wordt een consistente stijl gehanteerd: zo zijn de gegevens in het menu voor gedetailleerde informatie over een film altijd op dezelfde plekken. Dit bevordert gebruiksgemak. De visualisatie had wel iets breder gemogen. Bij het navigeren is het soms lastig om de juiste film weer terug te vinden.

Al met al, echter, is de visualisatie vrij goed. Alhoewel bepaalde designkeuzes zwak waren, slaagt het wel in het beoogde doel. Het geeft direct een weergave van goed en slecht beoordeelde films, en voor geïnteresseerden toont het bijzonder veel verzamelde data over de respectievelijk geselecteerde Hollywood film.

Comment on the color and interaction aspects (to Reading 4 and Reading 5).

Hoger kleuren contrast (critici: paars, publiek: roze). Het is beter als deze kleuren verder uit elkaar liggen. Qua interactie is het duidelijk zichtbaar om welke film het gaat tijdens het navigeren, alleen zitten de verschillende films vaak dicht op elkaar. Dit zorgt ervoor dat het lastig is om dezelfde film weer terug te vinden.

Suggest improvements and refinements.

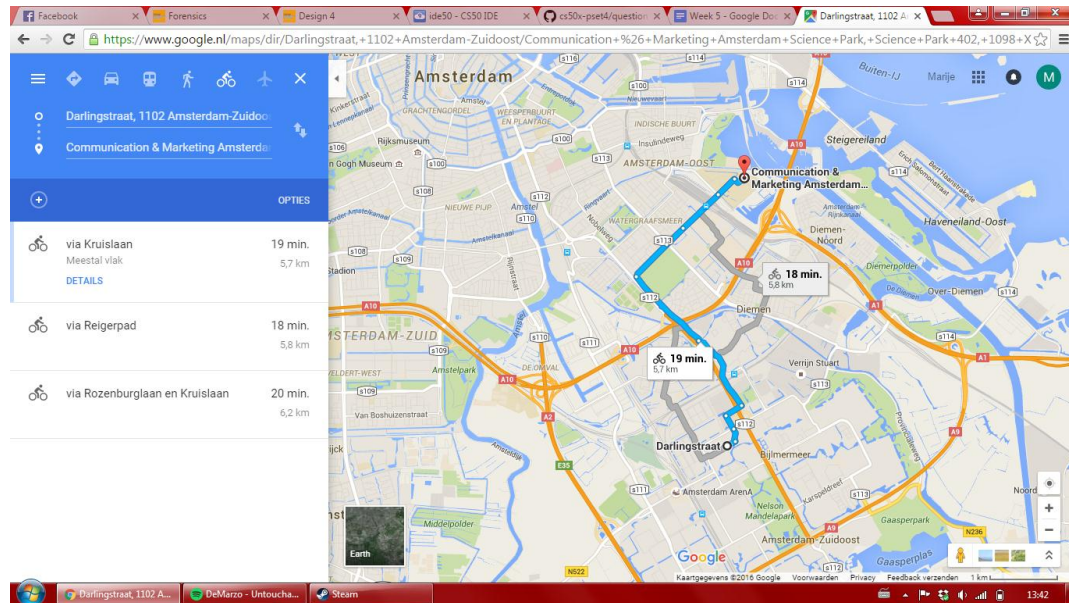
- Hoger kleuren contrast (critici: paars, publiek: roze). Het is beter als deze kleuren verder uit elkaar liggen.
- De iconen (cirkels) konden beter gekozen worden. Bijvoorbeeld een logo/karakter van de respectievelijke film.
- Herkomst van informatie (over budget, winst, enz.) aangeven. Zodat je de bronnen kan achterhalen en valideren. Enkel Metacritic voor de beoordelingen is gegeven.
- Labelen van de assen. Het is op het eerste oog niet duidelijk waarvoor de y-as en de x-as staan. De x-as wordt niet eens beschreven.

Problem 2: Patterns and colors in maps

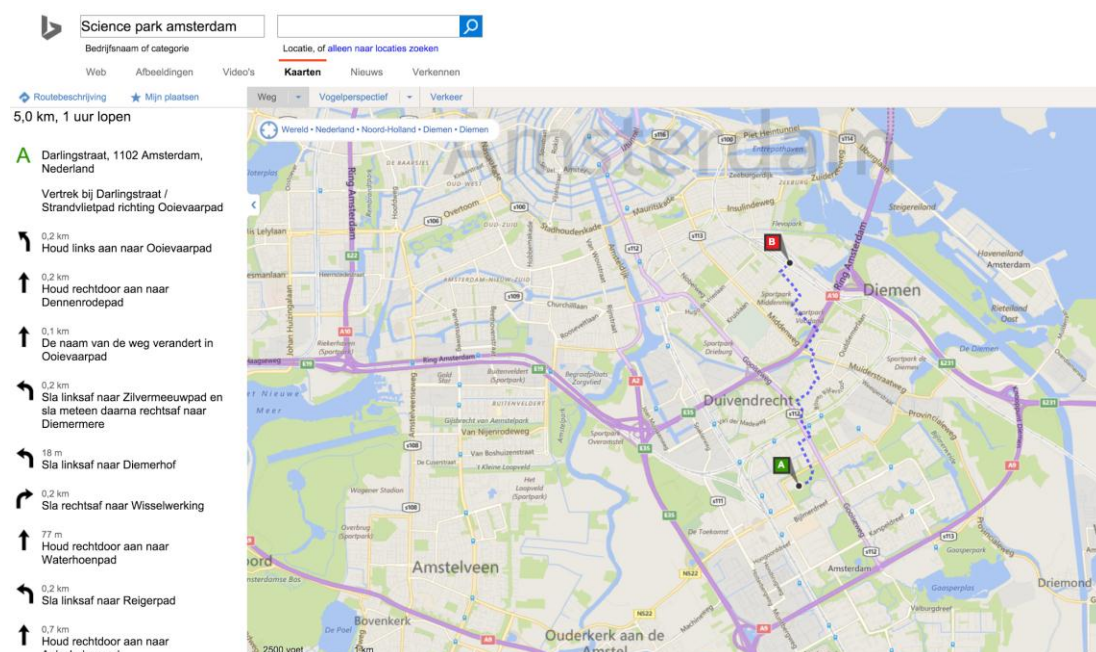
Patterns and colors are essential to maps. Compare a search for Harvard University on two interactive maps (e.g., Google Maps, Bing Maps, Yahoo! Maps, Apple Maps, map.harvard.edu). Answer the following questions, making references to concepts explained in Ware such as pattern recognition and properties of color. Please include screenshots of the examples you are comparing.

We hebben Google -en Bings Maps gebruikt voor het onderzoek.

screenshot van Google Maps:



Screenshot van Bing Maps:



Which map promotes an easier visual search for buildings?

Google maps. Want deze kaart is gedetailleerder. Hierdoor kan je je makkelijker oriënteren. Daarnaast staan er op Google Maps veel meer gebouwen of instellingen in vergelijking met Bing Maps. Ook heeft google maps satelite view en google streetview, beide heeft Bing Maps niet.

Which map more effectively visualizes routes from a random point A to point B?

Google maps. Deze geeft opties met routes. Zodat de gebruiker zelf een route kan kiezen waar hij/zij het meeste comfortabel mee is. Daarnaast heeft Google Maps meer opties voor vervoer, zoals fietsen, lopen, vliegtuig, auto, OV. Met Google Maps kan je het moment en de route van het OV kiezen, bij Bing kan dit niet.

Which map is an overall better visualization, and why?

Google Maps vonden wij beter dan Bing Maps. Dit omdat Google Maps een veel gebruikt platform is en ze hebben veel gegevens. Hierdoor heeft de kaart meer diepgang en beter zoekresultaten. Daarnaast zijn wij gewend om Google Maps te gebruiken en hierdoor vinden wij dit platform fijner om mee te werken. Google Maps biedt ook meer opties dan Bing.

Problem 3:

Find a multiple linked views visualisation on the web (see [Reading 5](#) for reference). Please include a screenshot and link of the visualization. Briefly summarize its intended objective and audience. Discuss all the interactive elements and their role in the chosen visualisation.

<http://ncva.itn.liu.se/research?l=en>



Figure: OECD eXplorer with three selected time-linked views map, scatter and fisheye bar chart for Europe TL3 regions.. 3 regions are highlighted in all views. The colour map shows "population age group 65+". <http://www.oecd.org/gov/regional/statisticsindicators/explorer>

<http://ncva.itn.liu.se/research/1.636660/OECDexplorerV4TL3EuropeageingpopulationNOSTORY.jpg>

Het OECD heeft dit onderzoek uitgevoerd om regionale ontwikkelingen beter in kaart te krijgen. Deze gegevens van de OECD helpt regeringen om beleidsmaatregelen bij te sturen, maar ook is het van wetenschappelijke interesse doordat de gegevens een beeld vormen van de invloed van politieke maatregelen op het welzijn. Regionale verschillen en ontwikkelingspatronen helpt de OECD om succesvol beleidsstrategieën te herkennen. Een hoog aandeel 65 plussers is doorgaans immers een teken van hoge mate van welzijn en welvaart. De doelgroep is dus kortom OECD experts, beleidsexperts, beleidsmakers, regionale planologen en journalisten.

De visualisatie geeft op drie verschillende wijzen aan hoe groot het aandeel 65+-ers per Europees land is. Dit wordt in exacte percentage weergegeven in de linker afbeelding. Donkerrood staat hierbij voor een relatief hoog percentage, terwijl lichtblauw staat voor een percentage naderend aan 0. De scatterplot rechtsboven is het aandeel 65 plussers op de x-as afgezet tegen het populatieaandeel 0-14 jarigen per land. Rechtsonder is een bar chart met eveneens exacte percentage per land, van hoog naar laag weergegeven. Dit zijn enkel de 24 landen met het hoogste percentage 65 plussers. Elk van deze drie visualisaties kan echter vervangen worden met een andere visualisatie met het menu boven de respectievelijke afbeelding. Dit is het interactieve aspect van de gehele multiple linked views visualisatie. Zodoende kan gekozen worden voor bepaalde dataweergaves / grafieken die tegelijk waargenomen kunnen worden. Deze weergaven zijn niet altijd hetzelfde; de scatterplot in de bovenstaande afbeelding geeft bijvoorbeeld andere data weer van hetzelfde fenomeen (populatieaandeel 65+). Ook kan bijvoorbeeld het populatieaandeel afgezet worden tegen de tijd in het venster rechtsonder. Omdat de data verschillende doelgroepen

aanspreekt, moet het flexibel en aanpasbaar zijn, zodat verschillende versies van vergelijkbare data gepresenteerd kunnen worden aan verschillende doelgroepen.