

Design Critique week 7

Data Processing

Yaleesa Borgman 6215262

December 2015

Intro, context en gebruik

De visualisatie betreft een visuele weergave van de arbeidsmarkt, hierin wordt aangegeven wat de verliezen en aanwinsten van banen zijn per sector over tijd. Deze is verdeeld in 2 visualisaties, een grafiek met een tabel en een grafiek met een lijngrafiek.

Deze visualisaties staan in de Wall Street Journal, dit is wel een blad voor goed opgeleide mensen. Ook is het publiek van de WSJ geïnteresseerd in de markten, dus zo ook de arbeidsmarkt en de trends hierin. Met een visualisatie kan je een trend goed zichtbaar maken. Wat dus de interesse zal wekken bij de lezers.

De bovenste visualisatie geeft in een interactieve grafiek aan in hoe veel procenten een arbeidsmarkt is verminderd of verbeterd per sector en ten opzichte van de maand ervoor. Zodra er met de muis wordt “ge-hovered” over een stipje die staat voor een sector, kan de gebruiker meer details zien. Ook verschijnen er zwarte stipjes in alle andere stipjes van dezelfde sector in de gehele grafiek, dit kan handig zijn voor het zien van trends. Ook zie je wanneer er een crisis is geweest om meteen te zien waarom er zoveel rood is in die periode.

In het tabel onder de eerste visualisatie staan de sectors nog een keer aangegeven maar nu op volgorde van grootste verandering in de sector (en alleen van november). Ook zie je hierbij de details en absolute aantallen, deze kunnen interactief ook op volgorde gezet worden.

De tweede visualisatie geeft op een effectieve manier heel veel data weer van over een grote periode van de jaren 50 tot nu. Zo kan er bijvoorbeeld meteen gezien worden dat er in de jaren 80 een dip zit en rond 2010. Onder deze grafiek bevindt zich nog een grafiek, deze geeft dezelfde data weer, maar dan in een lijngrafiek, om gemakkelijk de pieken en dalen te zien. Beide grafieken zijn aan te passen door een keuzemenuutje, hier kunnen verschillende etnische groepen, genders en leeftijden worden gekozen.

Design Principles

Er is vooral gebruik gemaakt van kleuren, grafieken en tabellen (visual encodings). Hierin zie je veel herhaling, de kleuren zijn divergent met een rood als negatief uiteinde en groen/blauw als positief uiteinde. De bolletjes die presenterend zijn voor de sectors komen elke maand in de grafiek weer terug, maar dan op een andere plek door de waardes die ze op dat moment hebben. Je kan de eerste grafiek ook zien als een hele dikke lijngrafiek, er kan een trend in worden gezien die vrij positief begint, gaat dan door een dipje om vervolgens weer te klimmen naar het positieve.

Verder zijn alle visualisaties vrij clean, er word vrij veel gebruik gemaakt van wit, en alleen kleur voor de data punten. Geen toevoeging van plaatjes of andere dingen die duiden dat het om de arbeid sector gaat behalve de titel.

Tufte's Principles

Scales

Bij de eerste visualisatie word gebruik gemaakt van een natuurlijke nullijn, boven de nullijn een positief gebied, onder de nullijn een negatief gebied. De data's staan op de x as, deze worden aangegeven in jaartallen. Omdat er tussen de jaartallen precies 12 bolletjes zitten die elk een maand voorstellen is dit een cleane oplossing.

Bij de tweede visualisatie is de x as de jaren en de y as de maanden. Deze worden opgedeeld in een kolommen/rijen grafiek waarbij elk vakje een maand moet voorstellen.

Lie factor

De lie factor heeft volgens tufte de volgende formule:

$$\text{Lie factor} = \text{size of effect shown in graphic} / \text{size of effect in data}$$

In de eerste visualisatie staat 1 bolletje gelijk aan 1 sector per 1 maand. Dit verandert niet in de gehele grafiek. De y as geeft het aantal sectoren weer in de negatieve of positieve kant, maar niet de data zelf, die is aangegeven met een kleur. Daar kan ik eerste oog opslag wat verwarring over zijn (maar uiteindelijk wel vrij duidelijk).

Data vs design variation

Wat wel opmerkelijk is, is dat de datapunten met 0.00% worden meegerekend in het positieve gedeelte, wat logischer is dan het negatieve gedeelte, maar als er bijvoorbeeld 5 sectoren die maand allemaal 0.00% hadden, lijkt het net of het

die maand erg positief ging in de arbeidsmarkt. Wat dus dan in werkelijkheid niet zo is.

Bij de tweede visualisatie is het duidelijk dat de kleuren duidelijk een 1% per stap in kleur is. ($2 + \text{aantal kleuren} = 11$)

Data ink ratio

Volgens tuft is er een formule om de data ink ratio uit te rekenen, oftewel of er onnodige elementen zijn gebruikt in de visualisatie, die de visualisatie niet nodig heeft om functioneel te zijn. De formule luid:

$$\text{Data ink ratio} = \text{data ink} / \text{total ink used in the graphic}$$

In de visualisaties word hij goed mee omgegaan. Ik kan niet echt iets bedenken wat weg gehaald kan worden.

Data density

$$\text{Data Density} = \text{number of entries in data array} / \text{area of data graphic}$$

Voor een 3 dimensionale dataset is veel ruimte nodig. Die op de manier an de eerste visualisatie vrij goed uitgebeeld word. Het is in essentie een grafiek met datapunten. De datapunten, hebben een plek, die niet direct gelinked is aan de plek waar deze staat, de echte data word directer aangegeven met kleuren. En dit allemaal in 1 grafiek, ik vind het knap.

De tweede visualisatie heeft 2 dimensionale data, maar kan een veel groter tijdsplan weergeven.

Layer information

Bij de dip in de eerste visualisatie staat in het lichtgrijs aangegeven dat dit een crisis periode is geweest. Wat meteen extra conclusie geeft aan de data, maar staat niet in de weg. Hij is vrij subtiel maar scheelt weer een extra denkstap.

Graphic design principles

Visual encodings

Eerste visualisatie:

Er word gebruik gemaakt van een grafiek, x as is tijd, y as is frequentie van sectoren. Verder word gebruik gemaakt van color value namelijk een diverging scale.

Tweede visualisatie:
Ook een grafiek, maar dan met duidelijke rijen en kolommen. Hier word ook gebruik gemaakt van een Diverging scale voor de color value.

Comments

Door het gebruik van zachte kleuren (ook die van de harde) ziet het er aantrekkelijk uit om te bekijken en mee te gaan spelen. Een leuke toevoeging bij de eerste visualisatie is dat als er over een data bolletje heen “ge-hovered” word, er zwarte puntjes verschijnen in dezelfde sectoren maar dan op een ander punt. Dit helpt je meteen om meer data in een keer te zien en op zoek te gaan naar trends.

Bij de tweede visualisatie is hetzelfde verhaal met hoveren, als er in de blokjes grafiek een data punt word bekeken, verschuift grafiek daaronder ook meteen van data punt en andersom. Erg leuke interactie.

Wat ik zelf nog anders zou doen is een zoekbalk voor de sectoren, iets van een drop-down menuutje. Je komt er wel vrij snel achter, zoveel verschillende sectoren zijn er niet. Maar het helpt je toch met zoeken. Ook zou ik toevoegen dat als je echt op een datapunt zou klikken, de rest wat vervaagd in contrast en de sectoren waar het om gaat naar voren springen (misschien met een lijntje erdoor) zodat je meteen een goede trend kan zien mocht de gebruiker daar in geïnteresseerd zijn.