



Wat data ons niet vertellen

De kracht van context en perspectief

Een van de interessantere scènes uit Adam McKay's *The Big Short* is wanneer Danny en Porter (gespeeld door Rafe Spall en Hamish Linklater) op onderzoek gaan in een nieuwbouwwijk in Florida om te kijken hoe waarschijnlijk het is dat de bewoners hun hypotheek kunnen betalen? Al snel komen zij erachter dat het merendeel van de huizen verlaten is en de bewoners die zij spreken geen benul hebben wat voor woekerhypotheek zij eigenlijk zijn aangegaan. Na deze inspectie in Florida, realiseren Danny en Porter zich dat de huizenmarkt inderdaad op instorten staat en zij shorten de markt tegen het advies van iedereen in.

Al vertelden de wiskundige modellen en de bijbehorende ratings op de hypotheekobligaties ons dat de Amerikaanse huizenmarkt zeer solide was, bleek dit in de werkelijkheid verre van waar, met de kredietcrisis van 2008 als gevolg. Deze crisis wordt vaak bestempeld als een black swan event, iets wat niemand zag aankomen. Maar hoe slaagden sommigen, zoals Danny en Porter, erin om dit wél te voorzien?

Dit had naar ons idee twee hoofdredenen. Ten eerste namen zij de data en de bijbehorende aannames over de huizenmarkt niet simpelweg aan als waarheid. Zij gingen actief toetsten of de aannames wel klopten. Ten tweede bekeken zij de situatie vanuit meerdere perspectieven. In plaats van alleen te vertrouwen op de ratings op hun computerscherm, reisden zij naar Florida en zagen met eigen ogen wat er werkelijk speelde.

J. Damen MSc (links) is Senior Consultant; F. Pardoel FRM MSc (midden) is Co-founder; en N. Cohen MSc is manager; allen werkzaam voor ADC Consulting.



INZICHT 1: VERTROUW DATA NIET BLINDELINGS

Laten we eerst inzoomen op de eerste reden: zij accepteerden de data over de huizenmarkt niet blindelings, maar stelden de vraag wat hun data daadwerkelijk maten en welke aannames eraan ten grondslag lagen.

De data en aannames waarop we onze modellen baseren, vormen slechts een vereenvoudigde weergave van de werkelijkheid. Net zoals een kaart een landschap representeert, zonder elk detail vast te leggen, zo ook geeft data ons slechts een beperkte inkijk in de werkelijkheid. Data tonen patronen en geven richting, maar vangen nooit alle nuances waarop diezelfde data zijn gebaseerd.

Een kaart kan ons helpen bij het navigeren door de bergen, maar het vertelt ons niet waar de losse stenen op het pad liggen en of de route ons gaat bevallen. Op dezelfde manier bieden data inzichten, maar missen zij vaak de context en subtiliteiten die essentieel zijn voor een volledig begrip van de situatie.

Een concreet voorbeeld hiervan is de koopgeschiedenis van iemand die boodschappen doet in de supermarkt. De kassabon laat zien dat deze persoon een gezonde maaltijd heeft gekocht, maar wat de bon niet vertelt, is dat hij na enige twijfel besloot die zak chips toch maar terug te leggen. De database registreert alleen wat er daadwerkelijk is afgerekend – niet het overwegingsproces dat eraan voorafging.

In dit geval zullen de gevolgen niet zo groot zijn: misschien ontvangt deze persoon volgende week minder relevante bonusaanbiedingen en loopt de supermarkt wat omzet mis. In andere situaties kunnen de verschillen tussen wat data ons vertellen en wat er werkelijk speelt, wel grotere gevolgen hebben. Neem bijvoorbeeld data die verzekeraars gebruiken:

- **Schadeclaims** voor nieuwe klanten en schadeverzekeringen voor klimaat- en natuurrampen worden vaak gebaseerd op historische data. De vraag is echter of deze data wel representatief en volledig zijn. Kortom, sluiten de omstandigheden van vroeger nog aan bij de risico's van nu en de toekomst?
- **Levensverzekeringspremies** worden onder meer bepaald door (zelfgerapporteerde) gegevens over gezondheid en leefstijl, zoals

gezondheidsklachten en rookgedrag. Ziekte of een ongezonde levensstijl kan leiden tot significant hogere premies of zelfs uitsluiting van de verzekering. Een grote verzekeraar stelt in zijn gezondheidsformulier: 'Wij gaan ervan uit dat u de vragen eerlijk, juist en volledig invult'. Gezien de prikkels is de vraag in hoeverre men daar echt vanuit kan gaan.

Een mogelijke oplossing is om niet alleen data te gebruiken die snel, goedkoop en makkelijk voorhanden zijn, maar om te evalueren welke data écht nodig zijn om de vraag zo volledig mogelijk te beantwoorden. Hierbij ook de evaluatie of er dan voldoende capaciteit en budget beschikbaar is om deze dataset samen te stellen.

Door kritisch te kijken naar de onderliggende data voor het specifieke vraagstuk kan worden voorkomen dat er onjuiste of onvolledige conclusies worden getrokken.

INZICHT 2: HET LOONT DATA VANUIT MEERDERE PERSPECTIEVEN TE BEKIJKEN

De tweede reden waarom Danny en Porter in *The Big Short* de kredietcrisis zagen aankomen was dat zij hun data vanuit meerdere perspectieven bekeken, en niet alleen maar vanuit een kwantitatieve hoek.

Een kwantitatieve aanpak werkt goed voor onderwerpen als natuurkunde en scheikunde. Maar deze aanpak schiet snel tekort zodra menselijk gedrag verklaard dient te worden. Of zoals beroemd natuurkundige Neil deGrasse Tyson ooit zei: 'In science, when human behavior enters the equation, things go nonlinear. That's why physics is easy and sociology is hard'.

Christian Madsbjerg biedt ons een potentiële oplossing voor dit probleem. In zijn boek *Sensemaking* duikt hij in het probleem dat wij als samenleving te gefocust zijn geworden op data en dat wij tegenwoordig de wereld alleen nog maar door de kwantitatieve lens willen begrijpen. Hij pleit voor het toevoegen van meer kwalitatieve data wanneer menselijk gedrag deel van je data beslaat. Ga bijvoorbeeld naar de plek waar de data worden verzameld en observeer

en interview mensen die de data generen. Ofwel, precies wat Danny en Porter hebben gedaan toen zij naar Florida afreisden.

Dit brengt ons bij een fundamenteel inzicht: *datasets zijn geen objectieve, wetenschappelijke feiten, maar sociale constructies*. Ze worden gevormd door de keuzes en aannames van degenen die ze creëren, verzamelen en interpreteren. Het toevoegen van kwalitatieve data kan helpen om blinde vlekken en onbewuste vooroordelen in data te verminderen.

Daarnaast kan ook een interdisciplinaire aanpak bijdragen aan het verbreden van het perspectief. Zet bijvoorbeeld data-analisten, verzekeraars en beleidsmakers bij elkaar wanneer er een nieuw model moeten worden ontwikkeld. Een data-analist kijkt waarschijnlijk anders naar een dataset dan een beleidsmaker. Door verschillende expertises te combineren ontstaat er een breder perspectief en wordt de kans kleiner dat cruciale nuances over het hoofd worden gezien.

Ook een scenario-gebaseerde aanpak, zoals beschreven door George Cairns in het boek *Scenario Thinking*, kan helpen. Hierbij brengen organisaties meerdere mogelijke toekomsten in kaart en testen ze strategieën onder verschillende omstandigheden. Dit vermindert de afhankelijkheid van één enkel model en vergroot de veerkracht, waardoor verkeerde aannames minder snel tot grote risico's leiden.

CONCLUSIE

De datarevolutie geeft ons ongekende mogelijkheden trends te herkennen, voorspellingen te doen en complexe vraagstukken te analyseren. Maar zoals de kredietcrisis van 2008 liet zien, kan blind vertrouwen op data en aannames zonder kritisch na te denken over wat de data werkelijk representeren gevaarlijk zijn.

Net als Danny en Porter moeten we verder kijken dan de cijfers op ons scherm. Door onze aannames te toetsen en verschillende perspectieven toe te passen, kunnen we betere beslissingen nemen en blijven we ons bewust van het feit dat data niet 'de waarheid' zijn, maar slechts een interpretatie daarvan. ■

Barometer van de pensioentransitie – hoe staan we ervoor?



AG JAARCONGRES 2025

dinsdag 3 juni
Spant! Bussum

Dagvoorzitter: Maarten van Meerten

Sprekers: dr. Eddy van Hijum | drs. ing. Ger Jaarsma | prof. dr. Bas Werker | prof. dr. Fieke van der Lecq

Paneldiscussie onder leiding van drs. Daan Kleinloog AAG

met onder andere Chantal de Groot MSc AAG | drs. Lieke Werner AAG | drs. Rob Schilder AAG