Wat is big data-visualisatie?

23-04-2015 16:32



Grote verzamelingen gegevens blijken meer informatie te bevatten dan op het eerste gezicht lijkt. Om die informatie er uit te krijgen, zijn wel slimme algoritmen nodig. Omdat de uitkomst van die algoritmen geen direct publiceerbare conclusie is, maar een nieuwe berg cijfers, moet de informatie alsnog toegankelijk gemaakt. Dat kan in woorden ('vooral vrouwen van middelbare leeftijd in middelgrote gemeenten lezen Heleen van Rooijen én rijden Daihatsu'), maar ook in grafieken. Het voordeel daarvan is dat de correlatie grotendeels automatisch zichtbaar kan worden gemaakt. Dat heet big data-visualisatie.

Het begrip big data wordt gebruikt om een (statistische) analyse aan te duiden, die wordt uitgevoerd over gegevens die zich niet in één database bevinden, maar over verschillende bronnen verspreid zijn. Juist door de combinatie van gegevens uit verschillende bronnen worden verbanden zichtbaar die niet uit een enkele bron te halen zijn. Die is immers met een bepaald doel samengesteld en dat doel blijkt uit de data: in een database met gegevens over boekverkopen staat niets anders dan dat. Een analyse van die database geeft alleen informatie over leesgedrag. Door die database te combineren met andere gegevens, bijvoorbeeld demografische gegevens, is een completer beeld van de doelgroep te construeren.