Data Processing – Design

David van Schie & Brent Kok

In onze Design opdracht willen wij aan het licht brengen dat de huidige Nederlandse huizenmarkt risicoverschijnselen laat zien, die grote gelijkenis vertonen met de huizenmarkt zoals deze was vlak voorafgaand aan de crisis ontstaan in 2008. Het belang, de oorzaak en de mogelijke gevolgen hiervan zullen wij op de volgende wijze aantonen. Allereerst zullen wij aantonen dat de Nederlandse economie sterk in de greep is van haar huizenmarkt. Hierbij zullen wij een aantal opmerkelijke feiten aanhalen die momenteel zorgen voor een oververhitting van de huizenmarkt:

* De verlaging van de overdrachtsbelasting naar 2% werkt oververhitting van de huizenmarkt juist in de hand, aangezien dit de kosten van de overdracht verlaagt);
* De lage rentestand (heeft tevens een negatief effect op de huizenmarkt aangezien dit oververhitting in de hand speelt).

Een ander punt dat zal worden aangehaald is:

* De relatief veel hogere *Loan-to-value* ratio in Nederland in vergelijking met onze buurlanden (dit zorgt ervoor dat mensen bij een afnemende economische groei eerder onder water zullen staan).

Bij elk van de bovenstaande punten zullen wij omschrijven wat het effect hiervan is op de huizenmarkt en het risicogehalte hiervan bij een omslag naar laagconjunctuur.

Wij zullen dit onderzoek uitvoeren met behulp van data afkomstig van de CBS, zoals de huidige huizenprijzen, het aantal verkochte woningen en recente wetswijzigingen. Hierbij zullen wij met name focussen op de overeenkomsten tussen de jaren 2008 en 2017. Een voorbeeld van relevante data: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/81885ned/table?ts=1519397763613

**Data types:** We zullen ons vooral richten op het data type *attribute.* Dit zijn meetbare eigenschappen, zoals de huizenprijs en de rentestand. Met deze waarden zullen we vergelijkingen door de tijd maken.  
  
**Dataset type:** De datasets die we gaan gebruiken zullen het meest weghebben van *tables*, aangezien er 2 waardes belangrijk zijn: een datum en een *attribute.*

**Dataset availability:** Onze data zal *static* of *offline* zijn. De data staat namelijk vast en is niet dynamisch. Er zullen geen nieuwe items worden toegevoegd, en waardes zullen ook niet aangepast worden.  
  
**Attribute types:** We maken gebruiker van *ordered data,* er is een duidelijke volgorde aanwezig in onze data, namelijk de datum. De data is ook *quantitative,* immers zijn waardes als huizenprijzen en rentestanden reële getallen waarmee gerekend kan worden en welke vergeleken kunnen worden.