



# Prijsindex Nieuwbouw Koopwoningen

Methodebeschrijving

projectnummer

28 maart 2024

# Inhoudsopgave

<b>1.</b>	<b>Inleiding</b>	<b>4</b>
1.1	Aanleiding	4
1.2	Doel PNK	4
1.3	Afkorting	4
1.4	Referenties	5
<b>2.</b>	<b>Berekeningsmethode PNK</b>	<b>6</b>
2.1	Indexberekening	6
2.2	Data	9
2.3	Berekening overige indicatoren	10
<b>3.</b>	<b>Berekeningsmethode aantal nieuwbouwtransacties</b>	<b>11</b>
3.1	Samenwerking CBS-Kadaster	11
3.2	Data	11
3.3	Methode afleiding nieuwbouwtransacties	11
3.4	Berekening aantallen en totale verkoopwaarde	11
3.5	Verkochte nieuwbouwoopwoningen t.o.v. opgeleverde nieuwbouwwoningen	12
	<b>Referenties</b>	<b>13</b>

# 1. Inleiding

## 1.1 Aanleiding

Het Centraal Bureau voor de Statistiek (CBS) berekent een Prijsindex voor Nieuwboukoopwoningen (PNK) naar aanleiding van een Europese verplichting. In Verordening 1148/2020 van de Europese Commissie [1] is vastgelegd dat lidstaten een Huizenprijsindex (HPI) moet berekenen. De verordening verwijst naar de Technical manual on Owner-Occupied Housing and House Price Indices [2] voor de eisen waaraan de HPI dient te voldoen. Onderdeel van de HPI is de PNK. Dit document beschrijft hoe het CBS de PNK berekent.

## 1.2 Doel PNK

De Prijsindex Nieuwbouw Koopwoningen (PNK) heeft als doel het meten van de prijsontwikkeling per kwartaal van verkochte nieuwbouwwoningen in Nederland. De woning moet gebouwd zijn op Nederlands grondgebied en moet verkocht zijn aan een particulier. Ook is het belangrijk dat de woning is of wordt opgericht op een nieuw fundament en voor de eerste keer wordt toegevoegd aan de woningvoorraad. Transformatiewoningen worden ook tot de populatie gerekend. Huur- en renovatiewoningen en woningsplitsingen behoren in dit geval niet tot de populatie. Verder geldt dat er geen sprake moet zijn van particulier opdrachtgeverschap (ook wel zelfbouw genoemd).

De PNK is opgenomen in de OOH (Owner Occupied Housing) en vanuit de richtlijnen van Eurostat voor OOH wordt een onderscheid gemaakt tussen nieuwbouw koopwoningen en zelfbouw. De PNK zoals die in OOH is opgenomen verschilt op één aspect van de originele PNK: in de wegingen van de deelreeksen van OOH is de grondwaarde eruit gehaald.

De PNK is daarnaast ook opgenomen in de HPI (House Price Index). Bij de HPI is de PNK inclusief de grondwaarde.

Gebruiksdoelen van PNK:

- Opname in de OOH. De OOH brengt de kostenontwikkelingen van huizenbezitters in kaart.
- Opname in de House Price Index (HPI). De HPI brengt de prijsontwikkelingen van nieuwbouw koopwoningen en bestaande koopwoningen (PBK) in kaart.
- Zelfstandige publicatie op Statline. De tabel wordt aangevuld met waarden en aantallen van nieuwbouwtransacties.

## 1.3 Afkortingen

De volgende afkortingen zijn in deze notitie gebruikt:

HDIF	Hedonische Double Imputation Fisher
HDIL	Hedonische Double Imputation Laspeyres
HDIP	Hedonische Double Imputation Paasche
HPI	House Price Index
OOH	Owner Occupied Housing
PNK	Prijsindex Nieuwbouw Koopwoningen

## 1.4 Referenties

[1] Europese Commissie (2020), verordening 1148/2020. Raadpleegbaar via <https://eur-lex.europa.eu/legal-content/NL/TXT/PDF/?uri=CELEX:32020R1148>

[2] European Commission – Eurostat (2017) Technical manual on Owner-Occupied Housing and House Price Indices. <http://ec.europa.eu/eurostat/documents/7590317/0/Technical-Manual-OOH-HPI-2017/>.

## 2. Berekeningsmethode PNK

### 2.1 Indexberekening

De prijsindex nieuwboukoopwoningen is een kettingindex en wordt daarnaast gewogen geaggregeerd. De PNK is opgebouwd uit twee deelreeksen naar regio. De eerste regio betreft vier grote steden van Nederland: Amsterdam, Den Haag, Rotterdam en Utrecht. De tweede regio betreft de rest van Nederland. Hiermee wordt gecorrigeerd voor uitschieters in bepaalde regio's. De berekening bestaat uit de volgende stappen, waarbij eerst de deelreeksen worden berekend en vervolgens de hoofdreeks voor Nederland totaal.

#### 2.1.1 Deelreeksen berekenen: vier grote steden en de rest van Nederland

##### Stap 1: korte indexreeksen berekenen

Om de PNK te kunnen berekenen worden in eerste instantie korte reeksen per jaar berekend. De PNK wordt berekend op basis van de hedonische regressiemethode. Dit is een meervoudige regressieanalyse, waarbij meerdere onafhankelijke variabelen van invloed zijn op de afhankelijke variabele. Voor de PNK betekent dit dat de samenhang wordt berekend tussen de verkoopprijs van de woning en verschillende woningkenmerken. Hierdoor kan de prijsontwikkeling gecorrigeerd worden voor kwaliteitsverschillen tussen de verkochte woningen in opeenvolgende perioden. Het volgende regressiemodel wordt gebruikt om een hedonische prijsindex te berekenen:

$$\log(p_i^t) = \beta_0^t + \beta_1^t x_{i1} + \dots + \beta_k^t x_{ik} + \varepsilon_i^t$$

In dit geval is  $\log(p_i^t)$  in periode  $t$  het natuurlijke logaritme van de verkoopprijs en  $x_{ik}$  is het  $k^e$  kenmerk van de  $i^e$  woning die verkocht is in periode  $t$ . Welke kenmerken in het model worden opgenomen is bepaald aan de hand van hun bijdrage aan het verklaren van de prijs. Alleen significant bevonden kenmerken zijn opgenomen.

Op basis van de coëfficiënten die de regressieanalyse oplevert, wordt een verkoopprijs geschat. De geschatte verkoopprijs is vervolgens input voor het berekenen van een prijsontwikkeling. In dit onderzoek is gekozen om de verkoopprijs van zowel de referentieperiode als van de verslagperiode te schatten. Deze wijze van hedonische regressie wordt ook wel de double imputation-methode genoemd [1].

De double imputation-methode kent drie varianten van hedonische regressie: Laspeyres, Paasche en Fisher. Alle drie zijn gebaseerd op dezelfde bovenstaande regressie, maar verschillen in hoe de prijsindex hier vervolgens uit wordt geconstrueerd. Gezien de resultaten van het onderzoek bleek de double imputation-methode van Fisher het meest geschikt, om een index voor Nederland totaal te berekenen. Deze index is het geometrische gemiddelde van de Laspeyres-index en de Paasche-index. De volgende formule is een weergave van de Hedonische Double Imputation Laspeyres Index ( $I_{HDIL}$ ):

$$I_{HDIL}(t,0) = \frac{\exp\left(\sum_{k=1}^K \beta_k^t \bar{x}_k^0\right)}{\exp\left(\sum_{k=1}^K \beta_k^0 \bar{x}_k^0\right)}$$

De volgende formule is een weergave van de Hedonische Double Imputation Paasche Index ( $I_{HDIP}$ ):

$$I_{HDIP}(t,0) = \frac{\exp\left(\sum_{k=1}^K \beta_k^t \bar{x}_k^t\right)}{\exp\left(\sum_{k=1}^K \beta_k^0 \bar{x}_k^0\right)}$$

De volgende formule is een weergave van de Hedonische Double Imputation Fisher Index ( $I_{HDIF}$ ):

$$I_{HDIF}(t,0) = (I_{HDIL}(t)I_{HDIP}(t))^{1/2}$$

In deze formules geeft  $\beta_k^t$  voor de geschatte coëfficiënt weer voor het  $k^e$  kenmerk in periode  $t$ ,  $\bar{x}_k^0$  staat voor het gemiddelde van het  $k^e$  kenmerk in periode 0,  $\bar{x}_k^t$  staat voor het gemiddelde van het  $k^e$  kenmerk in periode  $t$  en  $K$  geeft het aantal woningkenmerken weer dat is gebruikt in het hedonisch model.

De jaarreeksen worden berekend door voor elk kwartaal de ontwikkeling ten opzichte van het vierde kwartaal van het jaar ervoor te berekenen. In het eerste jaar wordt de ontwikkeling ten opzichte van het eerste kwartaal berekend. Er wordt gebruik gemaakt van onafgeronde PNK-cijfers. De eerste periode van de korte reeks wordt op 100 gezet.

Figuur 1. Korte reeksen

Korte reeksen	
2015 Q1	
2015 Q2	
2015 Q3	
2015 Q4	2015 Q4
	2016 Q1
	2016 Q2
	2016 Q3
	2016 Q4

Stap 2: Ontwikkeling ten opzichte van de vorige periode berekenen

Per periode wordt de ontwikkeling ten opzichte van de vorige periode berekend. De ontwikkeling wordt berekend door de waarde van het betreffende kwartaal te delen op de waarde van het voorgaande kwartaal.

(2)

$$\text{Ontwikkeling deelreeks\_PNK}^{q, q-1} = \frac{I_{\text{deelreeks\_PNK}}^q}{I_{\text{deelreeks\_PNK}}^{q-1}}$$

### Stap 3: Lange indexreeks berekenen

Met behulp van de berekende ontwikkeling per periode kan een index worden berekend voor de hele periode. De eerste periode van de reeks wordt op 100 gezet. De overige perioden worden berekend door de ontwikkeling te vermenigvuldigen met het indexcijfer van de vorige periode. Voorgaande wordt ook wel kettingen genoemd en het kettingen zorgt ervoor dat de invloed van de startperiode niet door blijft werken in de navolgende perioden.

### Stap 4: Herschalen

De index wordt herschaald naar het basisjaar. Hiervoor wordt de gemiddelde index van het basisjaar berekend. Vervolgens worden de indexcijfers van alle perioden gedeeld door dit gemiddelde en vermenigvuldigd met 100. Het basisjaar is 2020.

### Stap 5: Afronden

Tot slot worden de cijfers afgerond op één cijfer achter de komma.

## **2.1.2 Hoofdreeksen berekenen: Nederland totaal**

### Stap 1: Splits beide deelreeksen in korte indexreeksen

De jaarreeksen worden berekend door voor elk kwartaal de ontwikkeling ten opzichte van het vierde kwartaal van het jaar ervoor te berekenen. In het eerste jaar wordt de ontwikkeling ten opzichte van het eerste kwartaal berekend. Er wordt gebruik gemaakt van onafgeronde cijfers.

### Stap 2: Korte reeksen gewogen aggregeren

De korte jaarreeksen van de deelreeksen, dus voor vier grote steden en de rest van Nederland, worden gewogen geaggregeerd tot één korte jaarreeks per jaar. De totale verkoopwaarde per jaar dient als gewicht. De formule voor het gewogen aggregeren is:

(1)

$$I_{PNK\_NL\_totaal}^{j,q} = \frac{(W_{PNK\_wel\_g4}^j I_{PNK\_wel\_g4}^q + W_{PNK\_niet\_g4}^j I_{PNK\_niet\_g4}^q)}{(W_{PNK\_wel\_g4}^j + W_{PNK\_niet\_g4}^j)}$$

waarbij  $I_{PNK\_NL\_totaal}^{j,q}$  de index van de PNK voor Nederland totaal in kwartaal q van jaar j is,  $W_{PNK\_wel\_g4}^j$  de weging in jaar j van de PNK voor de vier grote steden en  $I_{PNK\_wel\_g4}^q$  de index in kwartaal q van de PNK voor de vier grote steden.

### Stap 3: Ontwikkeling ten opzichte van de vorige periode berekenen

Per periode wordt de ontwikkeling ten opzichte van de vorige periode berekend. De ontwikkeling wordt berekend door de waarde van het betreffende kwartaal te delen op de waarde van het voorgaande kwartaal.

(2)

$$\text{Ontwikkeling PNK\_NL\_totaal}^{q, q-1} = \frac{I_{PNK\_NL\_totaal}^q}{I_{PNK\_NL\_totaal}^{q-1}}$$



#### Stap 4: Lange indexreeks berekenen

Met behulp van de berekende ontwikkeling per periode kan een index worden berekend voor de hele periode. De eerste periode van de reeks wordt op 100 gezet. De overige perioden worden berekend door de ontwikkeling te vermenigvuldigen met het indexcijfer van de vorige periode. Voorgaande wordt ook wel kettingen genoemd en het kettingen zorgt ervoor dat de invloed van de startperiode niet door blijft werken in de navolgende perioden.

#### Stap 5: Herschalen

De index wordt herschaald naar het basisjaar. Hiervoor wordt de gemiddelde index van het basisjaar berekend. Vervolgens worden de indexcijfers van alle perioden gedeeld door dit gemiddelde en vermenigvuldigd met 100. Het basisjaar is 2020.

#### Stap 6: Afronden

Tot slot worden de cijfers afgerond op één cijfer achter de komma.

## 2.2 Data

### 2.2.1 Primaire waarneming

De data om de Prijsindex voor Nieuwbouw Koopwoningen te bereken is afkomstig uit een primaire waarneming. Dit houdt in dat in elk kwartaal middels een enquête gegevens over individuele woningverkoop worden gevraagd aan projectontwikkelaars en ontwikkelende bouwondernemingen. Er is een aselechte steekproef getrokken uit een geschatte populatie en deze wordt getrokken per stratum (regio X woningtype) en wordt regelmatig (gedeeltelijk) ververst.

### 2.2.2 Variabelen

Om de PNK te kunnen berekenen wordt gebruik gemaakt van garantiecertificaatgegevens (om de woningen te identificeren), de verkoopprijs en verschillende woningkenmerken.

#### *Verkoopprijs (afhankelijke variabele)*

Onder de verkoopprijs wordt het volgende verstaan: de overeengekomen (totaal)prijs inclusief grondkosten, kosten om de woning op naam te krijgen en eventueel andere bijkomende kosten, inclusief BTW. Indien bekend op het moment van verkoop, worden ook de kosten voor meerwerk meegenomen in de verkoopprijs.

#### *Verkoopperiode (variabele om te stratificeren)*

Het verkoopmoment is de datum waarop de (koop)overeenkomst door de laatste partij wordt ondertekend. Er is gekozen om het verkoopmoment aan te houden en niet het oplevermoment, omdat de prijs wordt bepaald en akkoord bevonden op het moment van tekenen. De oplevertermijn van een woning is daarnaast variabel per project en veelal ook per woning binnen een project. De tijd tussen het verkoopmoment en het oplevermoment kan daarbij behoorlijk oplopen in tegenstelling tot bij bestaande koopwoningen, waar deze termijn redelijk stabiel is.

De volgende woningkenmerken worden gebruikt om de PNK te bereken:

#### *Kenmerk 1: woonoppervlakte (onafhankelijke variabele)*

Onder woonoppervlakte wordt verstaan: de gebruiksoppervlakte van de woning volgens NEN 2580, bestaande uit de gebruiksoppervlakte van de woonfunctie aangevuld met de gebruiksoppervlakte van overige inpandige ruimte. De woonoppervlakte wordt opgegeven in vierkante meters.

#### *Kenmerk 2: woningtype (onafhankelijke variabele)*

Eén van de onafhankelijke variabelen is woningtype. Voor de deelreeks voor de vier grote steden wordt onderscheid gemaakt tussen eengezinswoning en meergezinswoning. Voor de deelreeks voor de rest van Nederland worden de volgende vijf woningtypen onderscheiden:

- Hoekwoning
- Tussen- of geschakelde woning
- Twee-onder-een-kap woning
- Vrijstaande woning
- Appartement

#### *Kenmerk 3: locatie (onafhankelijke variabele)*

Ondernemingen geven op het enquêteformulier aan wat de woonplaats is van de woning en de woonplaats is gebruikt om de provincie uit af te leiden. De provincienaam is vervolgens gebruikt in de berekening van de prijsindex.

#### *Kenmerk 4: garage (onafhankelijke variabele)*

Ondernemingen geven per woning aan of er een garage aanwezig is bij de woning en of deze garage inpandig of uitpandig is. Afhankelijk hiervan is de woonoppervlakte inclusief of exclusief de oppervlakte van de garage.

De overige uitgevraagde variabelen worden gebruikt ter validatie van de uitkomsten.

## **2.3 Berekening overige indicatoren**

Naast het indexcijfer van de PNK worden ook drie gerelateerde indicatoren berekend en gepubliceerd. Het betreft de indicatoren aantal verkopen, totale verkoopwaarde en gemiddelde verkoopwaarde. Deze indicatoren worden hieronder kort toegelicht.

### **2.3.1 Aantal verkopen**

Het aantal transacties van nieuwbouw koopwoningen is gebaseerd op data van het Kadaster.

### **2.3.2 Totale verkoopwaarde**

De totale verkoopwaarde van verkochte nieuwbouwwoningen wordt per deelreeks berekend door het aantal verkopen zoals geregistreerd door het Kadaster te vermenigvuldigen met de gemiddelde verkoopwaarde van een woning gemeten in de primaire waarneming.

### **2.3.3 Gemiddelde verkoopwaarde**

De gemiddelde verkoopwaarde van nieuwbouwwoningen wordt berekend door de totale verkoopwaarde (zie 2.3.2) te delen door het aantal verkopen (zie 2.3.1).

## 3. Berekeningsmethode aantal nieuwbouwtransacties

### 3.1 Samenwerking CBS-Kadaster

De prijsindex nieuwboukoopwoningen (PNK) maakt gebruik van data van verkopende ondernemingen. Omdat deze waarneming gebaseerd is op een steekproef, worden binnen dit onderzoek niet alle transacties van nieuwboukoopwoningen geregistreerd. Het Kadaster registreert wel alle woningtransacties. Uit deze registratie is echter niet direct af te leiden wanneer het een nieuwbouwwoning betreft. Voor het in kaart brengen van het aantal verkochte nieuwbouwwoningen is daarom in samenwerking met het Kadaster een aparte methode ontwikkeld om deze transacties te lokaliseren. Deze methode wordt hieronder toegelicht.

### 3.2 Data

De basis voor de methode is een combinatie tussen de Basisregistratie Kadaster (BRK) en de Basisregistratie Adressen en Gebouwen (BAG). Het transactiebestand van het Kadaster, opgebouwd vanuit de BRK, bevat alle woningtransacties in Nederland vanaf 1993. De database is gevuld aan de hand van informatie die is opgenomen in de overdrachtsakten.<sup>1</sup> De BAG bevat informatie over alle gebouwen in Nederland. Waaronder het gebruiksdoel van het object, het bouwjaar en de status van het pand (bijvoorbeeld bouwvergunning verleend, bouw gestart, in gebruik, etc.).

### 3.3 Methode afleiding nieuwbouwtransacties

De indicatie 'nieuwbouw' is bij het Kadaster geen authentiek gegeven, daarom is er een methodiek bedacht om de nieuwbouwtransacties te identificeren. Het is een stapsgewijze methode waarbij op basis van transactiedata van het Kadaster gecombineerd met data uit de BAG de nieuwboukoopwoningen worden geïdentificeerd.

Bij het filteren worden de onderstaande vragen bij elke transactie gesteld:

- Is er een koppeling tussen de transactie en een woning in de BAG?
- Dateert het perceel van vóór of na de start van de registratie door het Kadaster in 1993?
- Is het de eerste transactie van het perceel aan een natuurlijk persoon?
- Is het de eerste transactie van het object (woning) aan een natuurlijk persoon?
- Hoe groot is het verschil tussen het bouwjaar en het moment van verkoop?
- Heeft de transactie bij het Kadaster de identificatie nieuwbouwgrond bestemd voor een woning gekregen?
- Was de woning voorheen een huurwoning?

Een combinatie van de antwoorden op deze vragen leidt tot de indicatie of het gaat om een transactie van een nieuwboukoopwoning of niet.

### 3.4 Berekening aantallen en totale verkoopwaarde

Nadat de nieuwboukoopwoningen zijn afgeleid uit de Kadasterregistratie, wordt de som van het aantal transacties per periode berekend. Dit vormt het aantal dat weergegeven wordt in de tabel op Statline. De aantallen worden vervolgens vermenigvuldigd met de gemiddelde verkoopprijzen die berekend worden vanuit de primaire waarneming (hoofdstuk 2) aan de hand van de onderstaande formule. Dit vormt de totale verkoopwaarde die wordt weergegeven in de tabel op Statline.

---

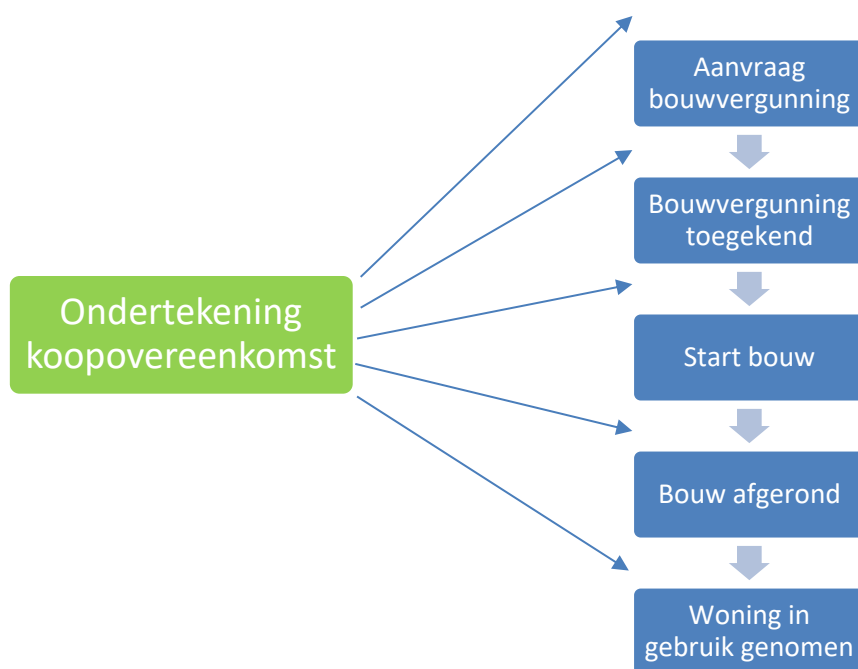
<sup>1</sup> De koopsom die het Kadaster registreert heeft betrekking op de grond, exclusief de bouwkosten van de nieuw te bouwen woningen. Om deze reden wordt de prijsindex niet gebaseerd op de Basisregistratie van het Kadaster.

$$P = N \cdot \bar{p}$$

Hierbij staat  $\bar{p}$  voor de gemiddelde verkoopprijs vanuit de primaire waarneming,  $N$  voor het berekende aantal uit de Kadasterregistratie en  $P$  voor de berekende totale verkoopwaarde.

### 3.5 Verkochte nieuwboukoopwoningen t.o.v. opgeleverde nieuwbouwwoningen

Het CBS publiceert naast aantallen van verkochte nieuwboukoopwoningen ook aantallen m.b.t. opgeleverde nieuwbouwwoningen<sup>2</sup>. Het voornaamste verschil tussen beide statistieken is het meetmoment: bij de opgeleverde nieuwbouwwoningen wordt een nieuwbouwwoning meegeteld op het moment dat de bouw van een huis is afgerond en wordt toegevoegd aan de woningvoorraad. Bij de verkochte nieuwboukoopwoningen wordt een woning meegeteld zodra er een koopovereenkomst is getekend. Zoals in onderstaande figuur is weergegeven, kan het ondertekenen van de koopovereenkomst plaatsvinden op meerdere momenten van het bouwproces.



Daarnaast worden bij de aantallen voor de PNK alleen koopwoningen meegenomen. De statistiek 'opgeleverde nieuwbouwwoningen' bevat naast koopwoningen ook huurwoningen.

<sup>2</sup> Zie: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/81955NED/table?dl=E4E5>

## Referenties

[1] Eurostat (2013). Handbook on Residential Property Prices Indices (RPPIs).

[2] Eurostat (2012). Technical manual on Owner-Occupied Housing for Harmonised Index of Consumer Prices (version 2.0).