

DREAM

Danish Research Institute for  
Economic Analysis and Modelling

DREAM

# Prisen på ejerboliger 1992-2021

En husstandsbaseeret undersøgelse af den geografiske  
variation i boligprisudviklingen

Jonas Zangenberg Hansen og Andreas Østergaard Iversen

August 2023



## **Prisen på ejerboliger 1992-2021**

©DREAM

August 2023

Rapporten er udarbejdet for  
Boligøkonomisk Videncenter

Forfattere:

Jonas Zangenberg Hansen

Andreas Østergaard Iversen

Rapporten er udgivet til digital brug og kan  
hentes på [www.dreamgruppen.dk](http://www.dreamgruppen.dk)

ISBN web: 978-87-91638-14-5

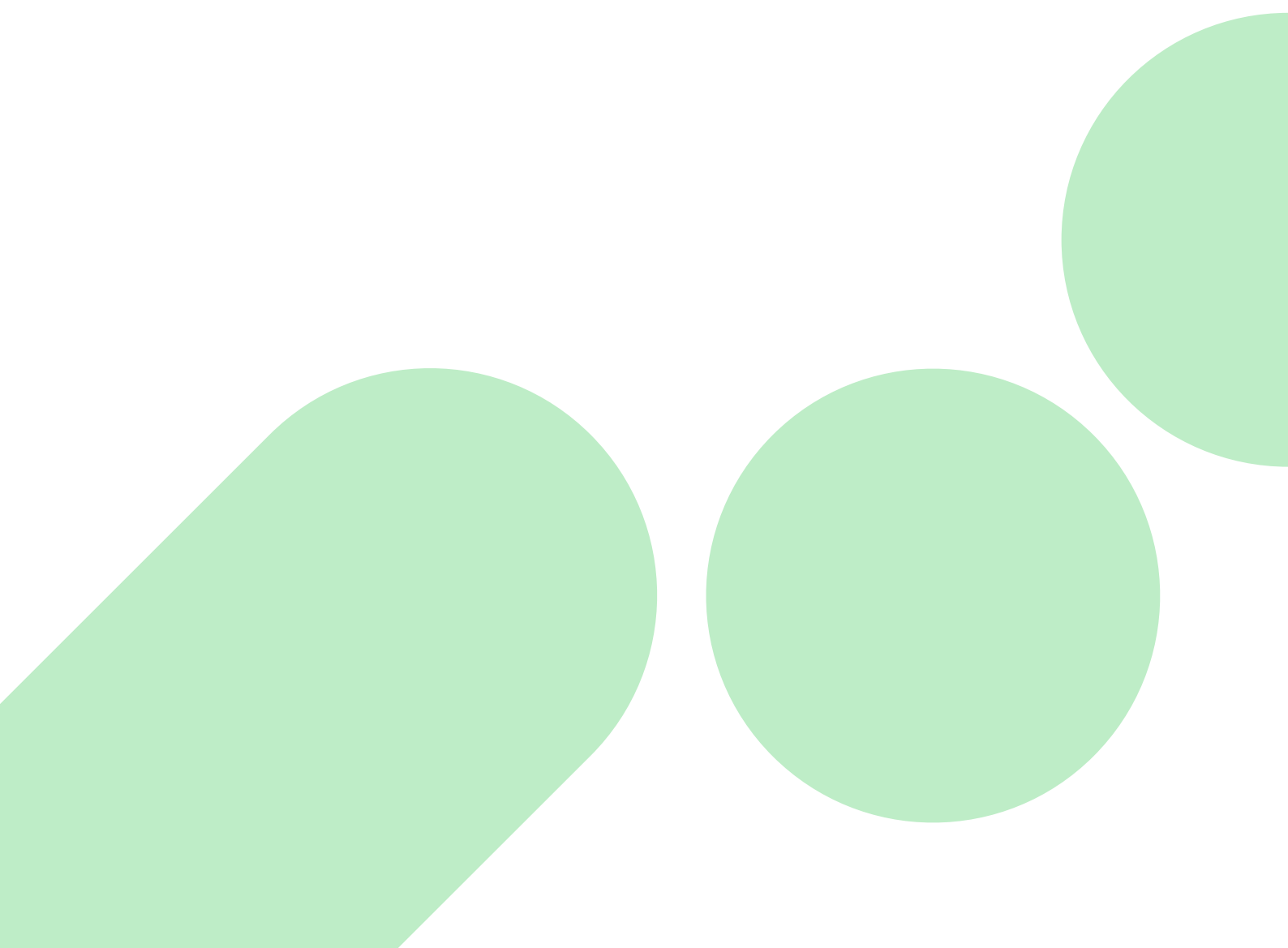
ISBN tryk: 978-87-91638-15-2

# Prisen på ejerboliger 1992-2021

En husstandsbaseeret undersøgelse af den geografiske variation i boligprisudviklingen

Jonas Zangenberg Hansen og Andreas Østergaard Iversen

August 2023





# Forord

Rapporten omhandler den historiske prisudvikling på ejerboliger i Danmark. For hver enkelt ejerbolig vurderes ejendomsværdien i perioden 1992-2021.

DREAM har tidligere foretaget lignende vurderinger af boligprisudviklingen. I juni 2015 blev der for Skatteministeriet udviklet en model til fremskrivning af ejendomsværdien af parcelhuse i historisk tid.

I januar 2018 udgav DREAM en undersøgelse af, hvorvidt boligejerne i perioden efter årtusindeskiftet har haft en økonomisk gevinst eller et tab ved at bo i ejerbolig frem for i en lejebolig af tilsvarende kvalitet og beliggenhed. Herunder vurderes størrelsen af boligejernes kapitalgevinst (eller tab), hvilket baseres på en historisk analyse af boligprisudviklingen på 25 pct. af ejerboligerne i Danmark. Undersøgelsen var finansieret af Boligøkonomisk Videncenter.

Fælles for DREAMs hidtidige undersøgelser er, at de baseres på geografisk vægtet regression (ofte kaldet GWR). Dette er en moderne beregningstung metode, der via sin geografiske vægtning giver en fin balance mellem lokale og globale egenskaber i beskrivelsen af et system.

I nærværende rapport præsenteres en ny vurdering af prisudviklingen på ejerboliger. I forhold til tidligere undersøgelser foretages en række korrektioner. Dels udvides den betragtede periode, ligesom stikprøvestørrelsen forøges til hele populationen af ejerboliger (100 pct.). Dels anvendes en ny og mere præcis metode til fastsættelse af ejendomsværdier.

Metoden anvendt i nærværende rapport baseres fortsat på, at der udføres en selvstændig vurdering af prisen på hver enkelt ejerbolig. Vurderingen sker ud fra en geografisk vægtning af handlede boliger i naboområdet til den betragtede bolig. Herudfra søger man at bestemme den lokale sammenhæng mellem boligens kvadratmeterpris og en række forklarende karakteristika på boligen.

Tidligere blev lokal lineær regression anvendt til at fastsætte kvadratmeterprisen. Denne regression er nu erstattet af en moderne metode baseret på maskinlæring (eng: machine learning). Herved opnås en mere præcis vurdering af boligprisudviklingen.

Rapporten er bestilt og finansieret af Boligøkonomisk Videncenter. Undersøgelsen er udført af DREAM, som har specificeret alle forudsætninger og beregningsmetoder anvendt i rapporten.

København, august 2023

# Resumé

Fra 1950 til begyndelsen af 1990'erne er prisudviklingen på ejerboliger nogenlunde ens på tværs af landet. Dette ændrer sig over de seneste 30 år. Rapporten fokuserer på den geografiske variation i udviklingen i ejerboligprisen gennem perioden 1992-2021.

## Metode til prisvurdering af ejerboliger

For samtlige 1,4 mio. privatejede ejerboliger foretages en selvstændig vurdering af boligens kvadratmeterpris i hvert kvartal i perioden 1992-2021. Den anvendte metode har tre væsentlige karakteristika.

For det første konstrueres en unik model for hver enkelt ejerbolig. Herved kan priseffekten af de indeholdte boligkarakteristika variere efter boligens beliggenhed.

For det andet anvendes geografisk vægtning. Kvadratmeterprisen på en ejerbolig fastsættes ud fra de nærmeste 15.000 handlede boliger. Hver observation vægtes, så bolighandler tæt ved den betragtede bolig vægtes højt, og salg længere væk vægtes lavere. Vægtningen sikrer balance mellem lokale og globale effekter i fastsættelse af boligprisen.

For det tredje fastsættes sammenhængen mellem boligens karakteristika og kvadratmeterprisen ved maskinlæringsalgoritmen LightGBM. Dette er en moderne metode, som kombinerer mange beslutningstræer (ved såkaldt boosting), ligesom algoritmen har kort afviklingstid og kan håndtere kategoriske variable.

Resultatet er en ganske præcis vurdering af boligpriserne, jf. Tabel R.1. At anvende maskinlæring øger den såkaldte PM20 med 2,7 pct. point i forhold til, hvis samme model estimeres med traditionel lineær regression.

**Tabel R.1**

### Træfsikkerhed ved diverse fejlmål

	Alle ejerboliger	Parcel- og rækkehuse	Ejerlejligheder
Gennemsnitlig afvigelse (MAE)	1.732 kr.	1.730 kr.	1.739 kr.
Gennemsnitlig procentvis afvigelse (MAPE)	16,5 pct.	18,2 pct.	11,3 pct.
Andel som rammer indenfor +/- 20 pct. (PM20)	76,1 pct.	72,4 pct.	87,6 pct.

Anm.: Træfsikkerheden måles for handlede ejerboliger. Her sammenholdes den faktiske handelspris med modellens prædiktionspris, idet prædiktionen foretages med boligens og boligejerens faktiske karakteristika på handelstidspunktet.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

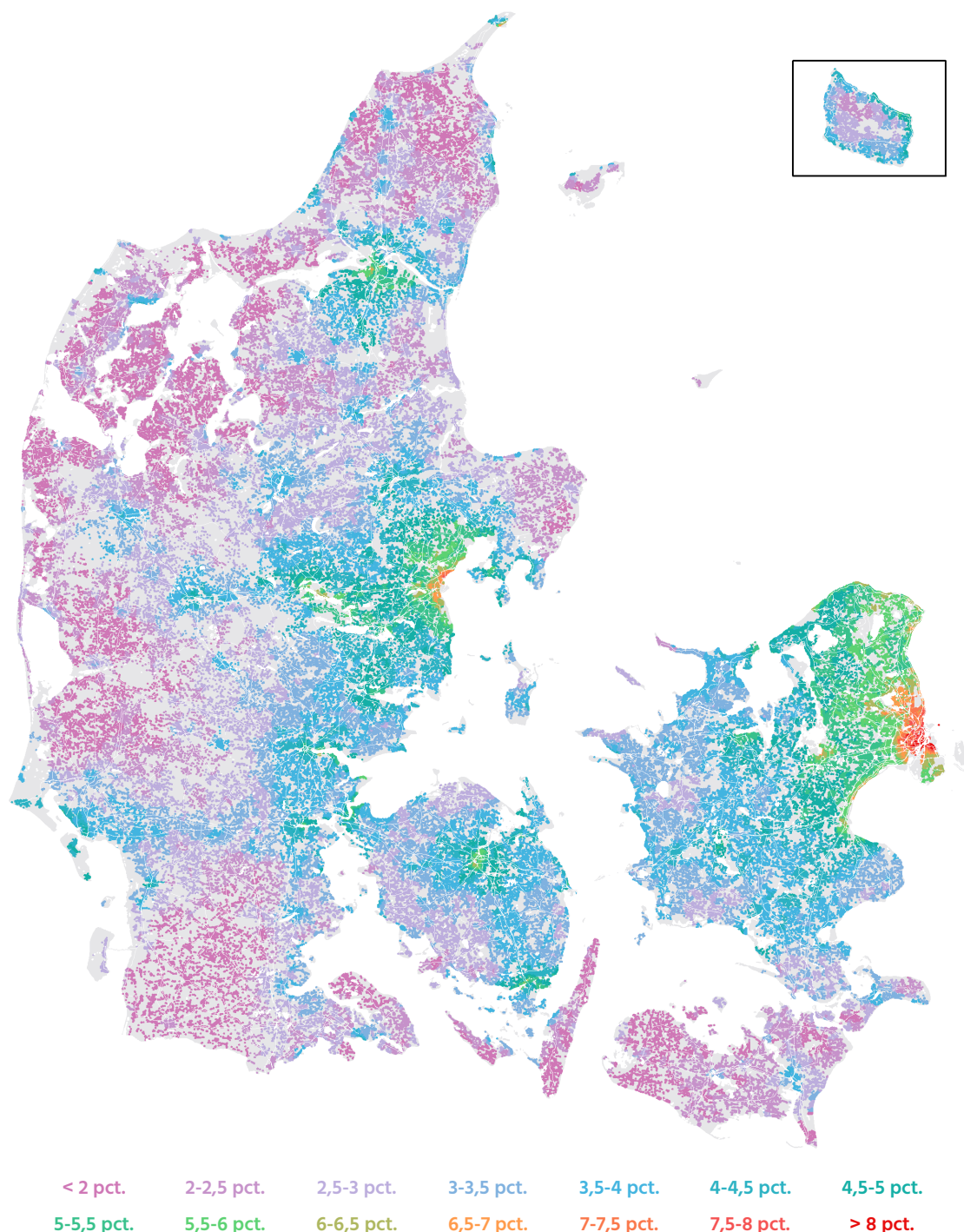
## Geografisk variation i prisudviklingen på ejerboliger

I hovedstadsområdet og i de større provinsbyer har boligpriserne i perioden 1992-2021 udvist en kraftigt stigende tendens. I de mindre bysamfund og i landdistrikterne har prisudviklingen i samme periode været mere afdæmpet, jf. Figur R.1 på den følgende side.

Prisstigningen er særligt stor i det centrale København. Herfra aftager prisstigningstakten efterhånden, som afstanden til bycentrum øges. En tilsvarende tendens ses omkring de lokale bycentre i de øvrige dele af landet. Blandt andre det centrale Aarhus, Odense, Aalborg og Esbjerg/Trekantsområdet er sådanne centre, hvor boligprisstigningen i perioden er kraftigere end i byernes opland, der igen stiger mere end landdistrikterne.

Figur R.1

Gennemsnitlig årlig ændring i boligprisen pr. kvadratmeter, 1992-2021



Anm.: Figuren viser den estimerede gennemsnitlige årlige procentvise ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Ejerboligerne inddeles i grupper efter deres skønnede kvadratmeterpris. Ejerboliger med de ti pct. laveste kvadratmeterpriser er geografisk spredt ud over det meste af Danmark fraregnet Sjælland og de større provinsbyer. I løbet af perioden 1992-2021 ses en tendens til, at de

billigste ejerboliger forsvinder i oplandet til de større byer. Modsat sker der i løbet af perioden en større koncentration af ejerboliger med en relativt lav kvadratmeterpris i de tyndere befolkede områder.

Ejeboligerne med de højeste kvadratmeterpriser er geografisk sammenklynget omkring Aarhus, København og langs kysten i Nordsjælland. Der ses en tendens til, at de dyreste ejerboliger målt på kvadratmeterprisen i løbet af de seneste 30 år sammenklumpes yderligere omkring centrum af København og Aarhus.



# Indhold

<b>1.</b>	<b>Indledning .....</b>	<b>10</b>
<b>2.</b>	<b>Introduktion til boligmarkedet og ejerboligen.....</b>	<b>13</b>
2.1	Boligformer i Danmark.....	13
2.2	Udvælgelse af ejerboliger, som indgår i undersøgelsen .....	16
2.3	Karakteristik af ejerboliger og deres ejere.....	17
<b>3.</b>	<b>Udvikling i ejerboligprisen .....</b>	<b>21</b>
3.1	Skøn over udvikling i boligpris 1992-2021.....	22
3.2	Regional udvikling i boligpriserne.....	24
3.3	Variation i boligprisen .....	33
<b>4.</b>	<b>Geografisk variation i ejerboligpriser .....</b>	<b>39</b>
4.1	Identifikation af dyre og billige områder.....	40
4.2	Geografisk variation i årlig vækstrate i boligpriserne .....	47
<b>5.</b>	<b>Metode til vurdering af boligprisudviklingen.....</b>	<b>59</b>
5.1	Illustrativt eksempel på geografisk vægtet regression.....	59
5.2	Geografisk vægtet regression.....	62
5.3	Sammenligning af modeller og valg af modeltype.....	65
5.4	Tuning af modellen.....	68
5.5	Estimationsresultater og træfsikkerhed .....	75
5.6	Forklarende variable .....	82
<b>6.</b>	<b>Referencer .....</b>	<b>85</b>
<b>A.</b>	<b>Ejerboliger efter prisdecil .....</b>	<b>87</b>
<b>B.</b>	<b>Årlig vækstrate i boligpriserne efter region.....</b>	<b>97</b>
<b>C.</b>	<b>Undersøgelsens datagrundlag .....</b>	<b>109</b>
C.1	Ejendomssalg .....	109
C.2	Sammenligning med øvrige boligprisindeks .....	110
C.3	Familieækvivalering af indkomst.....	113
C.4	Bystørrelseskode, landsdele og kommunegruppe .....	114

# 1. Indledning

Ejerboligprisen afhænger af boligudbuddet og boligefterspørgsel. Efterspørgslen påvirkes af en række samfundsøkonomiske forhold såsom renteniveau, adfærdsændringer og urbanisering. Udbudssiden er præget af knaphed og træghed. Prisudviklingen på ejerboliger har stor betydning for den finansielle stabilitet og for den makroøkonomiske udvikling.

Ejerboliger handles på markedsvilkår, og prisen fastsættes som et resultat af en forhandling mellem køber og sælger. Prisniveauet bestemmes i udgangspunktet af udbud og efterspørgsel.

På overordnet niveau påvirkes boligefterspørgslen af samfundsøkonomiske forhold såsom renteniveau, lønstigningstakt, ledighedsgrad, den generelle prisstigningstakt og den demografiske udvikling som eksempelvis befolkningstilvækst og aldersfordeling. For eksempel kan store årgange af ældre øge efterspørgslen efter ejerlejligheder og rækkehuse på bekostning af parcelhuse. Regler som fastsætter eksempelvis belåningsmuligheder og værdien af rentefradrag spiller ligeledes ind på boligpriserne på nationalt niveau.

En række faktorer influerer boligefterspørgslen på regionalt niveau. Dette kan være forhold som urbanisering, ændrede erhvervsstrukturer eller kulturelle faktorer som for eksempel adfærdsændringer. For eksempel kan flere singler øge efterspørgslen efter mindre boliger beliggende i de større byer.

Lokalt er forhold såsom beliggenhed og vedligeholdelsesstand afgørende for boligefterspørgslen og prisdannelsen. Ejerboliger beliggende tæt ved en befærdet vej eller ved en togbane må alt andet lige forventes at have en lavere efterspørgsel, hvilket påvirker prisen i nedadgående retning. Det modsatte gør sig gældende for ejerboliger beliggende kystnært.

Boligudbudssiden er præget af knaphed og træghed, hvilket kan betyde at udbuddet ikke nødvendigvis kan matche for eksempel en stigning i boligefterspørgslen. Trægheden i boligudbuddet dannes, da det kan tage lang tid at opføre nye boliger. På boligmarkedet kan knaphed opstå i attraktive områder, hvor der er begrænset byggekapaletitet, mangel på nye byggegrunde eller reguleringsmæssige hindringer.

Når boligefterspørgslen eksempelvis er højere end udbuddet, vil ejerboligprisen øges til et højere niveau. En gruppe af potentielle købere har nu ikke længere mulighed for at købe, hvorved efterspørgslen dæmpes. Når boligpriserne stiger på grund af knaphed, bliver det trods trægheden mere fordelagtigt at opføre nye boliger, hvorved boligudbuddet hvis muligt øges.

Markedsmekanismen sikrer at markedet justerer sig over tid, og der etableres ligevægt mellem udbud og efterspørgsel.

## **Boligpriserne påvirker samfundsøkonomien**

Boligprisudviklingen har stor betydning for samfundet. For de fleste boligejere er boligen den største enkeltstående investering, vedkommende foretager i sit liv. Prisudviklingen er afgørende for husstandens formueudvikling.

Stigende huspriser stimulerer privatforbruget og øger incitamentet til at bygge nyt. Stigende priser vil øge forskellen mellem en boligs markedspris og gælden i boligen. Denne såkaldte

friværddi kan belånes, hvilket blandt andet kan påvirke privatforbruget og bygge- og vedligeholdelsesaktiviteter. Hvis ikke belåningen øges, vil kapitalgevinsten blive realiseret, når boligen sælges. Modsat kan faldende boligpriser dæmpe privatforbruget.

Udviklingen i boligprisen kan have effekt på uligheden i samfundet. Dels mellem ejere og lejere, og dels mellem boligejere i forskellige dele af landet. Siden begyndelsen af 1990'erne har boligprisen udviklet sig markant forskelligt på tværs af landet, hvor især de større byer har oplevet markante husprisstigninger.

Relativt høje boligpriser i visse geografiske områder gør det sværere for lavindkomstfamilier at købe bolig i området. Dette kan påvirke bosættelsesmønstret og mobiliteten mellem områder.

### **Bedre indblik i den geografiske spredning i boligprisudviklingen**

At opnå en bedre forståelse af den geografiske spredning i boligprisudviklingen er central, idet den gennemsnitlige udvikling sjældent er repræsentativ for mindre geografiske områder eller egne med særlige kendetegn.

I nærværende undersøgelse vurderes prisudviklingen for hver enkelt ejerbolig i landet i perioden 1992-2021. Dette giver et unikt indblik i den geografiske variation.

At have et kvalificeret bud på ejendomsværdien for hver bolig over en lang tidsperiode giver desuden en række enestående muligheder. Eksempelvis kan størrelsen af boligkapitalgevinster vurderes, og det kan betragtes, hvilke persongrupper og hvilke geografiske områder, der i særlig grad har fået fordel af de over en årrække stigende priser.

Desuden kan boligprisvurderingen anvendes til at beregne hver husstands belåningsgrad. I kombination med oplysninger om husstandens økonomiske situation kan dette anvendes til at vurdere, hvor udsat husstanden er overfor økonomiske udsving som for eksempel en rentestigning (ofte kaldet stresstest). Sådanne beregninger er væsentlige til at vurdere den finansielle stabilitet.



## 2. Introduktion til boligmarkedet og ejerboligen

I ejerboligen disponerer ejeren over sin bolig via et skøde, der tinglyses på ejendommen. Dette er den mest udbredte ejerform i Danmark. I undersøgelsen af prisudviklingen på ejerboliger fokuseres på privatejede enfamilie ejerboliger, som hovedsageligt anvendes til boligformål. Disse udgør omkring 83 pct. af samtlige ejerboliger.

En boligejer køber typisk sin ejerbolig, når vedkommende aldersmæssigt befinder sig mellem midt 20'erne og frem til slutningen af 30'erne. Dette sker i takt med, at man typisk finder en fast partner og stifter familie. I så fald efterspørges en mere permanent bolig, som gerne må være med udearealer tilknyttet. Mange flytter derfor mod parcel- og rækkehuse, som typisk er ejerboliger. Fra slutningen af 60'erne fraflyttes ejerboligen efterhånden, som boligtagerne ønsker en mindre bolig og en bolig med mindre vedligeholdelse end ejerboligen.

Ejerboliger ejes derfor typisk af personer i de arbejdsduelige aldre. Dette er medvirkende til, at boligejerne typisk har en relativt høj husstandsindkomst sammenlignet med befolkningen som helhed.

### 2.1 Boligformer i Danmark

Primo 2021 fandtes i omegnen af 3,1 mio. boliger i Danmark, jf. Tabel 2.1. Boligerne opdeles i to hovedgrupper. For det første omkring 2,9 mio. helårsboliger<sup>1</sup> i form af enfamiliehuse (stuehuse til landbrugsejendomme, parcelhuse og rækkehuse), etageboliger (lejligheder) og øvrige boliger, som typisk vil være kollegieboliger, døgninstitutioner eller en blanding af bolig og erhverv. Den anden hovedgruppe omfatter cirka 244.000 sommer- og fritidshuse. I modsætning til helårsboligen er det ikke tilladt at bo i et sommerhus hele året, hvis huset ligger i et område, der er udlagt til sommerhuse<sup>2</sup>.

Den mest udbredte boligform i Danmark er ejerboligen. Her disponerer ejeren over sin bolig via et skøde, der tinglyses på ejendommen. Dette står i modsætning til både andelsboligen, hvor beboeren har brugsret og andel i foreningens formue, men ikke ejer selve boligen, og lejeboligen, hvor beboeren kun har brugsret. Lejeboliger kan opdeles i tre overordnede typer efter udlejers ejerforhold: Almene boliger, private udlejningsboliger og offentlige udlejningsboliger.

Boks 2.1 indeholder en oversigt over de anvendte boligformer for helårsboliger.

<sup>1</sup> De fleste helårsboliger beliggende i boligområder og byer er behæftet med bopælspligt. Som udgangspunkt indebærer dette, at hvis en bolig ikke er beboet af ejeren, så skal boligen i stedet lejes ud således, at den er beboet. Det er kommunen, som beslutter, om en helårsbolig er behæftet med bopælspligt. Siden foråret 2013 har kommunerne haft mulighed for at give flexboligstatus til helårsboliger i landområder, så bopælspligten ophæves, og man dermed kan bruge boligerne som fritidshus.

<sup>2</sup> Der kan dog opnås dispensation, så sommerhuset kan benyttes til helårsbolig.

## Boks 2.1

### Definition af ejerformer

---

#### Helårsboliger

*Ejerboliger* afgrænses som ejerlejligheder og øvrige.

Ejerlejligheder identificeres som privatejede boliger (ejet af privatpersoner, interessentskaber, selskaber samt selvejende institutioner), der er beliggende i etageejendomme og udskilt i ejerlejligheder (dvs. boligen har et selvstændigt ejerlejlighedsnummer).

Øvrige ejerboliger end ejerlejligheder omfatter blandt andet parcelhuse, rækkehuse og øvrige enfamiliehuse, der ikke er ejet af andelsboligforeninger, almene boligorganisationer eller det offentlige.

En andel af ejerboligerne vil være udlejet, for eksempel ejerlejligheder, der er lejet ud. Disse udlejede ejerboliger indregnes som ejerboliger.

*Andelsboliger* opgøres som bolig- og erhvervsenheder, der ejes af private andelsboligforeninger. I tilgængeligt data er det ikke muligt at afgøre, hvorvidt andelsboligen er beboet af andels haveren, eller om andelsboligen er udlejet.

*Almene boliger* identificeres ved samtlige boliger beliggende i ejendomme ejet af en almen boligorganisation. Almene boliger opføres med offentlig støtte. Huslejen fastsættes, så den enkelte boligafdeling akkurat løber rundt (omkostningsbestemt leje), hvorfor den generelle stigning i boligpriserne ikke påvirker lejen. Nogle almene boliger er forbeholdt beboere med særlige behov. Kommunerne har anvisningsret til almene boliger.

*Private udlejningsboliger* er boliger ejet af privatpersoner, selskaber eller selvejende institutioner. Boligerne omfatter blandt andet egentlige udlejningsejendomme. Nogle gange indregnes udlejede ejerboliger – for eksempel ejerlejligheder, der lejes ud – som private udlejningsboliger. Dette er ikke tilfældet i nærværende undersøgelse.

*Offentlige udlejningsboliger* er boliger ejet af kommune, region eller stat og som udlejes til borgerne. Boligerne er typisk målrettet en bestemt type personer som for eksempel unge, handicappede eller ældre.

#### Sommer- og fritidshuse

Sommer- og fritidshuse omfatter enheder, hvis primære anvendelse er sommerhus, kolonihavehus og anden enhed til fritidsformål. Er der givet tilladelse til helårsbeboelse, omklassificeres boligen til i stedet at være et fritliggende enfamiliehus.

Ejerboliger udgjorde primo 2021 omkring 57 pct. af bestanden af helårsboliger i Danmark. En del ejerboliger bebos imidlertid ikke af ejeren selv men lejes i stedet ud. Fraregnes disse boliger bestanden af ejerboliger, udgør ejerboliger omkring halvdelen af den samlede boligbestand<sup>3</sup>.

---

<sup>3</sup> Andelen af ejerboliger i Danmark opgøres i visse andre undersøgelser til omkring halvdelen af boligbestanden. Dette gælder blandt andet Danmarks Statistik (2015). Forskellen i ejerboligers andel af den samlede boligbestand skyldes hovedsageligt, hvorvidt udlejede ejerboliger indregnes som ejerboliger eller private udlejningsboliger. Modsat for eksempel Danmarks Statistik indregnes udlejede ejerboliger i denne undersøgelse som netop ejerboliger, hvorfor ejerboligers andel af den samlede boligmasse øges. Følgelig indregnes en udlejningsbolig beboet af ejeren som udlejningsbolig, hvilket mindsker ejerboligers andel af boligbestanden, idet for eksempel Danmarks Statistik (2015) indregner sådanne boliger som ejerboliger.

Den næstmest udbredte ejerform er almene boliger, som udgør omkring hver femte bolig. Herefter følger private udlejningsboliger (knap 13 pct. af alle helårsboliger), andelsboliger (7,5 pct.), mens offentlige udlejningsboliger er den mindst udbredte ejerform (2 pct.). En del ejerboliger bebos ikke af ejeren selv, men disse lejes i stedet ud. Dette omfatter 15 pct. af ejerboligerne primo 2021. Knap 5 pct. af ejerboligerne – eller 4,5 pct. af alle enfamiliehuse – står uden persontilmelding. En del af disse boliger uden persontilmelding anvendes øjensynligt som fritidsboliger, og Møller (2017) skønner, at i størrelsesordenen 0,4-1,5 pct. af alle enfamiliehuse i Danmark er reelt forladte.

**Tabel 2.1**  
**Boliger efter ejerform, primo 2021**

	Enfamiliehuse	Etageboliger	Øvrige	I alt
<b>Helårsboliger</b>				
I alt	1.678	1.210	15	2.902
Ejerbolig	1.377	279	4	1.659
– Beboet af ejer	1.175	123	1	1.298
– Udlejet	135	136	1	273
– Ubeboet	67	20	1	88
Andelsbolig	55	150	0	205
Almen bolig	163	420	1	584
Privat udlejningsbolig	59	327	3	390
Offentlig udlejningsbolig	23	34	5	62
Uoplyst ejerform	0	0	1	1
<b>Sommer- og fritidshuse</b>				
I alt	-	-	244	244
– Beboet	-	-	23	23

Anm.: Antal boliger er opgjort i tusinde boliger. En bolig regnes som beboet, hvis mindst en person er tilmeldt adressen ifølge Det Centrale Personregister (CPR).

Kilde: Egne beregninger på baggrund af Danmarks Statistiks Boligopgørelse.

Boligmassens sammensætning efter ejerform varierer betydeligt på tværs af landet. Som en overordnet angivelse af en boligs beliggenhed anvendes bystørrelsen for ejendommens geografiske placering. Kilden er Danmarks Statistiks Byopgørelse, der angiver bystørrelsen for placeringen af samtlige boliger i Danmark, jf. Danmarks Statistik (2023).

Byområderne grupperes dels i det sammenhængende byområde hovedstadsområdet, som dækker et stort byområde omkring København. For områder uden for hovedstadsområdet opgøres bystørrelsen efter antal indbyggere i byområdet. Grupperingen af bystørrelse kan ses i Tabel 2.2, og grupperingen beskrives nærmere i appendiks C.2.

Ejerboligen er generelt mest udbredt i mindre befolkede områder. Jo større byområde, der betragtes, jo mindre andel af boligbestanden udgør ejerboligen. Fra at udgøre ni af ti boliger i landdistrikterne, falder ejerboligers andel af alle helårsboliger inden for et afgrænset område til lige over hver tredje bolig i hovedstadsområdet. I de større byområder stiger i stedet andelen af både almene og private udlejningsboliger. Hovedstadsområdet er specielt, idet dette er det eneste område i Danmark, hvor andelsboliger udgør mere end 5 pct. af den samlede boligbestand. I København udgør andelsboliger knap hver tredje bolig.

Tabel 2.2

Helårsboligers efter ejerform og bystørrelse, primo 2021

	Ejerformens andel af samtlige boliger i området, pct.					Antal Tusind boliger
	Ejer	Andel	Almen	Privat udlejning	Offentlig udlejning	
Hele landet	57	7	20	13	2	2.902
<b>Efter bystørrelse</b>						
Hovedstadsområdet	38	20	26	15	2	653
- København	31	30	18	20	1	378
Byer ≥ 100.000 indbyggere	44	4	29	21	2	302
- Aarhus	45	4	29	20	3	143
- Odense	48	4	27	20	2	93
- Aalborg	39	4	32	24	1	67
Byer 40.000-99.999 indbyggere	50	4	27	16	2	345
Byer 30.000-39.999 indbyggere	52	4	31	11	2	84
Byer 20.000-29.999 indbyggere	50	4	30	15	2	153
Byer 10.000-19.999 indbyggere	58	4	23	13	2	198
Byer 5.000-9.999 indbyggere	62	5	21	10	3	195
Byer 2.000-4.999 indbyggere	67	5	15	9	3	250
Byer 500-1.999 indbyggere	76	4	10	8	2	232
Byer < 500 indbyggere	90	0	2	7	1	352

Anm.: Boliger med uoplyst ejerform er udeladt af tabellen, men indgår i antal boliger i kolonnen yderst til højre. Bystørrelse for en boligs beliggenhed opgøres ud fra Danmarks Statistiks Byopgørelse. Boliger med uoplyst bystørrelse er udeladt af tabellen. København omfatter København og Frederiksberg kommuner. Aarhus, Odense og Aalborg er de samlede byområder.

Kilde: Egne beregninger på baggrund af Danmarks Statistiks Boligopgørelse og Byopgørelse.

## 2.2 Udvælgelse af ejerboliger, som indgår i undersøgelsen

Undersøgelsen baseres på data fra en række administrative registre, der indeholder oplysninger på befolkningen og boligerne i Danmark. De enkelte registre anvendt i undersøgelsen er nærmere beskrevet i appendiks A, herunder beskrives data for ejendomssalg.

I princippet burde undersøgelsen blive gennemført på samtlige 1,65 mio. ejerboliger i Danmark. Det er imidlertid nødvendigt at foretage en række datamæssige afgrænsninger, hvorved undersøgelsens datagrundlag begrænses til at omfatte omkring 1,4 mio. ejerboliger. Dette svarer til cirka 83 pct. af samtlige ejerboliger i landet.

Der foretages to væsentlige dataafgrænsninger: For det første betragtes kun privatejede ejerboliger, som hovedsageligt anvendes til beboelse. Herved ekskluderes 11,7 pct. af samtlige ejerboliger, jf. Tabel 2.3. Primært udelukkes ejerboliger, som ikke er privatejede (disse er i stedet primært ejet af selskaber), eller ejerboliger beliggende på en meget stor grund. Yderligere ekskluderes boliger, der optræder med et urealistisk stort boligareal i forhold til almindelig beboelse eller enhedens erhvervsareal udgør mere end halvdelen af det samlede areal.

For det andet er det nødvendigt at kunne finde en række personspecifikke karakteristika på boligejerne, idet disse karakteristika anvendes i undersøgelsen. Dette gælder eksempelvis ejernes indkomstforhold, der blandt andet anvendes til at skønne over boligens værdi. Som



følge heraf ekskluderes boliger, hvor ejerne ikke kan identificeres i tilgængeligt data. Sidstnævnte gruppe kan for eksempel være danskere, der (midlertidigt) opholder sig i udlandet. Samlet set udgår 5,5 pct. af det samlede antal ejerboliger, idet de nødvendige personspecifikke karakteristika ikke kan identificeres, eller boligens placering er ukendt.

**Tabel 2.3**

**Udvælgelse af privatejede ejerboliger til beboelse, primo 2021**

	Antal boliger, tusind	Andel af alle ejerboliger, pct.
Ejerboliger i alt	1.659	100,00
<b>Afgrænsning til privatejede beboelsesejendomme</b>		
Bolig større end 400 kvadratmeter	5	0,31
Erhvervsareal udgør mere end 50 pct.	2	0,13
Ikke kategoriseret som helårsbeboelse	< ½	0,01
Bolig ikke privatejet	138	8,29
Grund større end 10 hektar	49	2,96
<b>Manglende oplysninger i data</b>		
Ikke alle ejere registreret	< ½	0,002
Bolig ejet af mere end to ejere	5	0,30
Manglende oplysninger på mindst en ejer	39	2,37
Boligens beliggenhed ikke registreret	47	2,86
<b>Resultat af ejendomsudvælgelse (øverste række fratrukket øvrige rækker)</b>		
Ejerboliger udvalgt til undersøgelse	1.374	82,77

Kilde: Egne beregninger på baggrund af Danmarks Statistiks Boligopgørelse og Byopgørelse.

## 2.3 Karakteristik af ejerboliger og deres ejere

I det følgende gives en kort karakteristik af ejerboligen i forhold til beliggenhed og boligtype og udviklingen i ejerboligbestanden siden årtusindeskiftet.

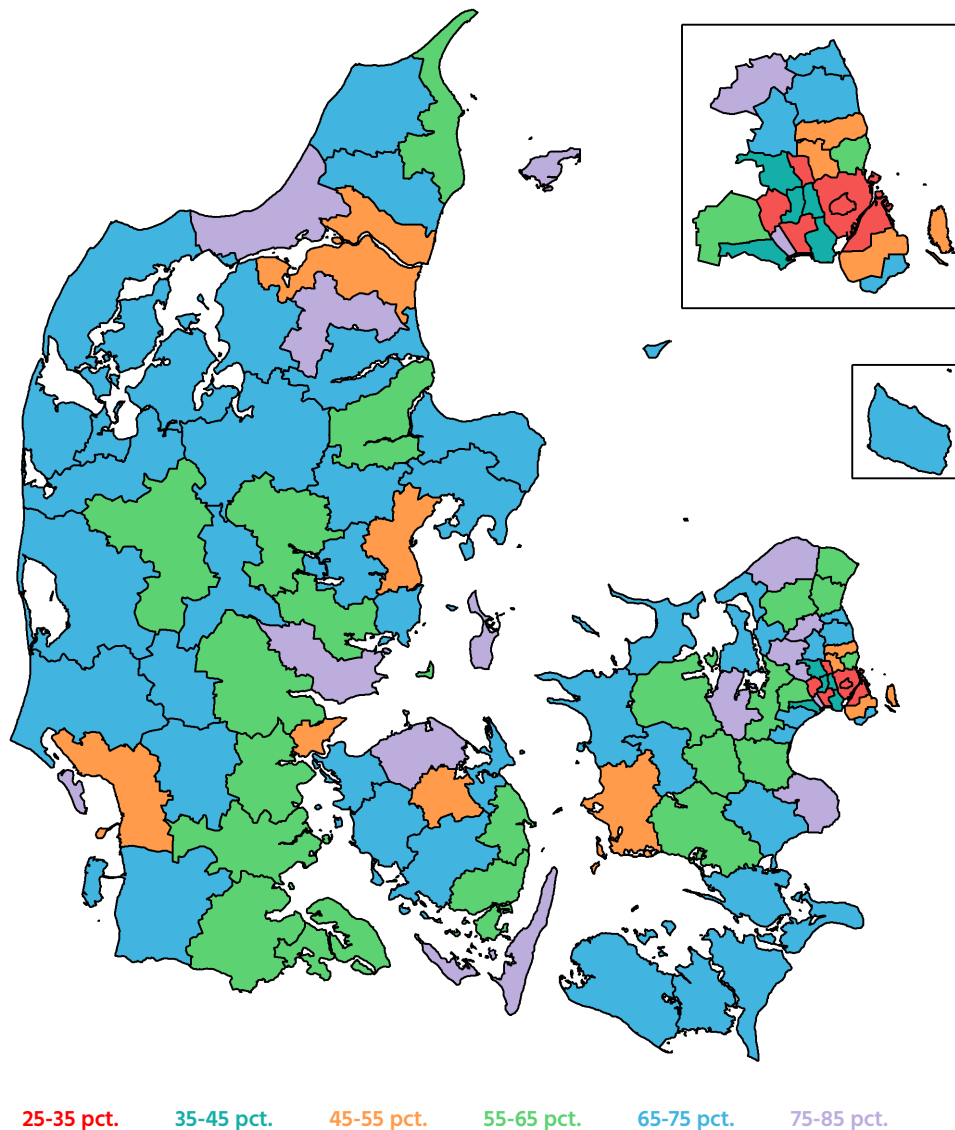
### Ejerboligers beliggenhed

Koncentrationen af ejerboliger til helårsbeboelse er generelt størst uden for de større byområder, jf. Figur 2.1. I de små ø-kommuner udgør ejerboliger mindst tre ud af fire boliger i kommunen. Langs den jyske vestkyst (dog fraregnet Esbjerg kommune), på store dele af Fyn samt på Lolland-Falster udgør ejerboliger 65-75 pct. af det samlede antal helårsboliger.

Ejerboligers andel af boligbestanden aftager, når større byområder betragtes. I Trekantsområdet samt på Midtsjælland udgør ejerboliger 55-65 pct. af alle helårsboliger i kommunerne. I de største bykommuner aftager ejerboligers andel af boligbestanden yderligere. I Odense og Aalborg kommuner udgør ejerboliger cirka halvdelen af alle boliger, mens andelen er 46 pct. i Aarhus. Ejerboliger udgør den mindste andel af boligbestanden i en række kommuner omkring København. Her udgør ejerboliger omkring hver tredje bolig.

**Figur 2.1**

Ejerboliger som andel af boligbestanden efter kommune, primo 2021



Anm.: I Vallensbæk, Læsø, Samsø og Ærø kommuner udgør ejerboliger den største andel af boligbestanden, idet andelen i disse kommuner overstiger 80 pct. I Københavns kommune udgør ejerboliger cirka 30 pct. af boligbestanden, hvilket er den mindste observerede andel.

Kilde: Egne beregninger på baggrund af Danmarks Statistiks Boligopgørelse.

### Ejerboligers type

Størstedelen af ejerboligerne er parcelhuse, idet disse udgør to tredjedele af samtlige ejerboliger, jf. Tabel 2.4.

**Tabel 2.4**

**Boligtipe for udvalgte ejerboliger, primo 2021**

	Boligtypens andel af samtlige ejerboliger i området, pct.				Antal Tusind boliger
	Stuehus	Parcelhus	Rækkehus mv.	Etagebolig	
Hele landet	6	66	11	17	1.659
<b>Efter bystørrelse</b>					
Hovedstadsområdet	<1	35	13	52	247
Byer $\geq$ 100.000 indbyggere	<1	45	16	39	134
Byer 40.000-99.999 indbyggere	<1	67	12	20	172
Byer 30.000-39.999 indbyggere	<1	57	20	23	44
Byer 20.000-29.999 indbyggere	<1	67	16	17	77
Byer 10.000-19.999 indbyggere	<1	72	16	12	115
Byer 5.000-9.999 indbyggere	<1	79	15	6	120
Byer 2.000-4.999 indbyggere	1	86	10	3	168
Byer 500-1.999 indbyggere	2	91	6	1	177
Byer < 500 indbyggere	29	70	1	<1	318

Kilde: Egne beregninger på baggrund af Danmarks Statistiks Boligopgørelse og Byopgørelse.



### 3. Udvikling i ejerboligprisen

Siden begyndelsen af 1900'erne ses en markant geografisk variation i boligprisudviklingen. Prisen er særligt steget i københavnsområdet og de øvrige større byer. Udviklingen har medført, at spredningen i boligprisen på tværs af landet er betydeligt større i dag, end det tidligere var tilfældet.

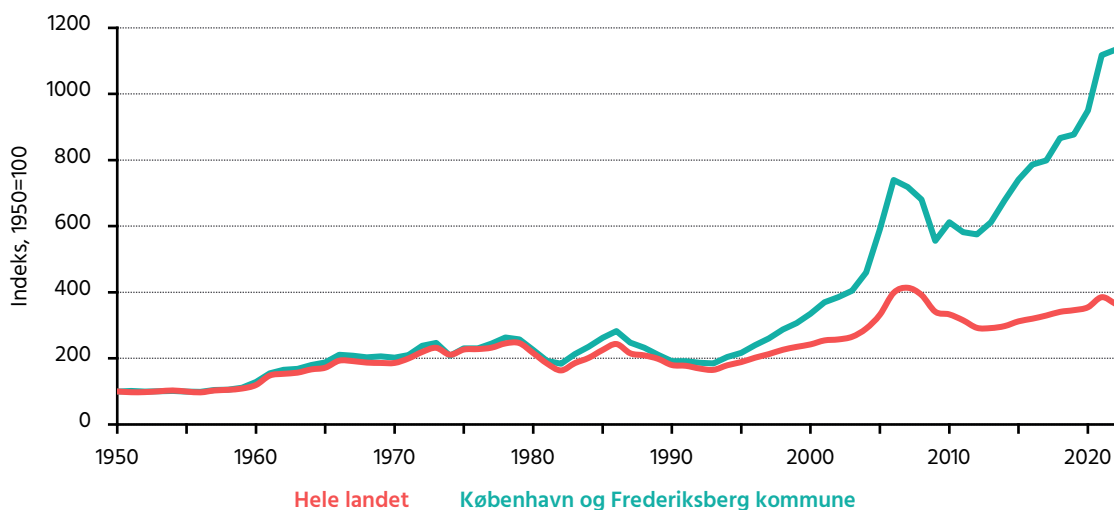
Ejerboligmarkedet har gennemgået en iøjnefaldende udvikling gennem de seneste årtier. I perioden fra 1950 til begyndelsen af 1990'erne var boligprisudviklingen nogenlunde ens på tværs af landet, jf. Figur 3.1. Over en 30-årig periode, begyndende midt i 1960'erne, følger prisstigningerne nogenlunde inflationen, dog med markante udsving undervejs i perioden.

I begyndelsen af 1980'erne falder boligpriserne som følge af stigning i både arbejdsløsheden og renterne. I 1986 kom kartoffelkuren, som reducerede værdien af rentefradraget, hvilket gjorde det dyrere at låne penge og dermed også dyrere at bo i ejerbolig. Da arbejdsløsheden samtidig fortsatte med at stige, blev resultatet faldende boligpriser, som først ophørte i begyndelsen af 1990'erne.

Over de sidste cirka 30 år er boligpriserne steget mere end den generelle prisudvikling. Stigningen har været markant forskellig på tværs af landet. I hovedstadsområdet og i de større provinsbyer har boligpriserne udvist en kraftigt stigende tendens. I de mindre bysamfund og i landdistrikterne har prisudviklingen i samme periode været mere afdæmpet.

I perioden siden begyndelsen af 1990'erne har boligefterspørgslen især været rettet mod de større byer, hvor der ofte kun er få ledige byggegrunde. Dette har drevet boligprisen i byerne op relativt til øvrige dele af landet. Der er flere årsager til, at boligefterspørgslen i særlig grad har rettet sig mod byerne.

**Figur 3.1**  
Real boligpris for enfamiliehuse 1950-2022



Anm.: Figuren viser boligprisudviklingen korrigeret for inflation.

Kilde: Historiske boligprisindeks fra Boligøkonomisk Videncenter.

I den resterende del af kapitlet fokuseres der på de store regionale forskelle i boligprisudviklingen fra begyndelsen af 1990'erne til i dag. Gennem perioden skønnes boligprisen for hver enkelt ejerbolig ved en statistisk model.

Når boligprisudviklingen beskrives, betragter man sædvanligvis, hvordan salgspriserne på boliger udviklede sig inden for et afgrænset geografisk område. Dette kan være problematisk i det omfang, de solgte ejendomme ikke var repræsentative for hele området. I et område med både nybyggeri og eksisterende boliger, kan det trække den observerede gennemsnitlige salgspris pr. kvadratmeter op, hvis en stor andel af salgene er nybyggeri, hvor kvadratmeterprisen er højere end i de eksisterende ejendomme. For at rense for denne skævhed, kan man vægte hver salgspris i forhold til, hvor stor en andel netop denne type bolig udgør af områdets samlede boligbestand.

I denne undersøgelse er valgt en tilgang, hvor der for hvert kvartal i perioden 1992-2021 laves et skøn over boligprisen for hver enkelt bolig<sup>4</sup>. Dette giver en unik mulighed for at illustrere boligprisudviklingen i forskellige dele af landet, herunder de områder, hvor der i perioder ikke handles mange ejendomme.

### 3.1 Skøn over udvikling i boligpris 1992-2021

I denne undersøgelse skønnes over prisudviklingen på samtlige ejerboliger i Danmark. Dette sker ud fra en metode, som vurderer hver boligs værdi på baggrund af prisen på solgte ejerboliger i nærområdet til den enkelte bolig.

Beskrivelse af den historiske boligprisudvikling har ofte startpunkt i 1992. Dette skyldes, at detaljerede salgsdata stammer fra tinglysningen af ejendomshandler, som digitaliseres i netop dette år. Det er værd at holde sig for øje, at boligpriserne i perioden 1986-1993 faldt realt med 36,5 pct., jf. Figur 3.1. Undersøgelsen starter således umiddelbart efter et relativt stort prisfald på ejerboligmarkedet, og den gennemsnitlige pristigningstakt ville have været anderledes, såfremt det var muligt at begynde i 1986.

#### Skøn over prisudvikling ved geografisk vægtet regression

For at skønne over prisen for hver enkelt ejendom i Danmark anvendes den statistiske metode geografisk vægtet regression. Ideen i metoden er, at man for hver enkelt ejerbolig fastsætter netop denne boligs handelspris på baggrund af kvadratmeterprisen på handlede boliger i lokalområdet til den betragtede bolig. Den gennemsnitlige kvadratmeterpris for området justeres i forhold til individuelle karakteristika ved den enkelte bolig. Det kan være, at boligen er beliggende i første række til kysten, hvilket gør den dyrere end naboejendommene i anden og tredje række. Boligen kan være beliggende lige op af en større vej, så prisen er lavere end i området på grund af støjgener. Boligen kan have et tag af asbestplader, der kan trække mod en lavere salgspris end lokalområdets gennemsnitspris.

På denne måde vurderes en individuel handelspris på alle ejerboliger i Danmark. Den skønnede pris kan forstås som den instinktive fornemmelse en ejendomsmægler har, når vedkommende skal vurdere en ejendom.

Metoden til at foretage disse skøn over boligpriser er kort beskrevet i Boks 3.1. I kapitel 5 gives en uddybende beskrivelse af estimationsmetoden. Metoden estimerer handelsværdien

---

<sup>4</sup> Andre metoder kan anvendes til at vægte handelspriserne på solgte ejendomme. Ved offentliggørelse af prisindeks for ejerboliger anvender for eksempel Danmarks Statistik den værdivægtede SPAR-metode (Sales Price Appraisal Ratio), jf. Danmarks Statistik (2023). Boligøkonomisk Videncenter anvender boliger handlet flere gange til at regne prisændringer ('repeat sales'-metoden), jf. Boligøkonomisk Videncenter (2023). Også i Boligmarkedsstatistikken vægtes de realiserede handler i forhold til den samlede boligbestand, jf. Boligmarkedsstatistikken (2023).

for samtlige 1,4 mio. privatejede enfamilie ejerboliger i Danmark i hvert kvartal i perioden 1992-2021.

For at vurdere salgsprisen for hver enkelt ejerbolig anvendes detaljerede handelsdata fra tinglysningsystemet. I systemet er der en vis forsinkelse, før alle salg registreres. Ved undersøgelsens afslutning var det seneste tilgængelige dataår med alle bolighandler registreret således 2021.

### Boks 3.1

#### Metode til fastsættelse af prisudvikling

For hvert enkelt ejerbolig i Danmark anvendes en avanceret model til at give et skøn over boligens værdi i hvert kvartal i perioden 1992-2021. Modellen anvender geografisk vægtning og en moderne teknik baseret på maskinlæring (eng: machine learning).

Metoden tager udgangspunkt i lokal regressionsanalyse. Ideen er, at der dannes én unik model for hver enkelt bolig, hvor boligprisen ønskes estimeret. Modellen bestemmer sammenhængen mellem boligens karakteristika og den lokale kvadratmeterpris.

Sammenhængen fastlægges ud fra et datamateriale bestående af alle solgte boliger i et cirkelformet område med den betragtede bolig som centrum. Området indeholder cirka 15.000 bolighandler og dannes unikt for hver enkelt bolig, hvor prisen skal vurderes. Hver handlet bolig tildeles en vægt i forhold til den geografiske afstand til den bolig, netop denne model vedrører. Handler tæt ved boligen selv tildeles en høj vægt, og boliger længere væk tildeles en lavere vægt.

Karakteristika så som boligens størrelse, opførelsesår, grundstørrelse mv. plus forklarende kendetegn ved boligens beliggenhed som for eksempel afstand til kyst, motorvej m.fl. har betydning for fastsættelse af kvadratmeterprisen. Desuden medtages karakteristika på ejerne af boligen, idet det formodes at for eksempel ejernes indkomstniveau udtrykker noget om boligens vedligeholdelsesmæssige stand.

I modellen bestemmes boligens prisudvikling over tid. Dette giver mulighed for at danne et unikt prisindeks for den enkelte bolig. Herved får man et skøn over, hvad den lokale udvikling i boligprisen har været på lige præcis den enkelte adresse.

Analysen udføres for alle boliger, uanset om boligen er solgt i løbet af den betragtede periode eller ej.

Træfsikkerheden måles på solgte boliger. Her sammenlignes boligens estimerede værdi med den faktiske handelspris. Metodens bedste træfsikkerhed opnås i Hovedstadsområdet, hvor 83 pct. af boligerne estimeres indenfor +/- 20 pct. af deres sande handelspris. Generelt er træfsikkerheden højest i de større byer, hvilket er forklaret af, at boligerne i større byer er mere homogene end boligerne på landet. Desuden er der typisk flere ejendomssalg i byområder. På landsplan er træfsikkerheden på 76,1 pct. (ejerlejligheder: 87,6 pct., enfamiliehuse: 72,4 pct.).

Den anvendte metode er grundigt forklaret i kapitel 5.

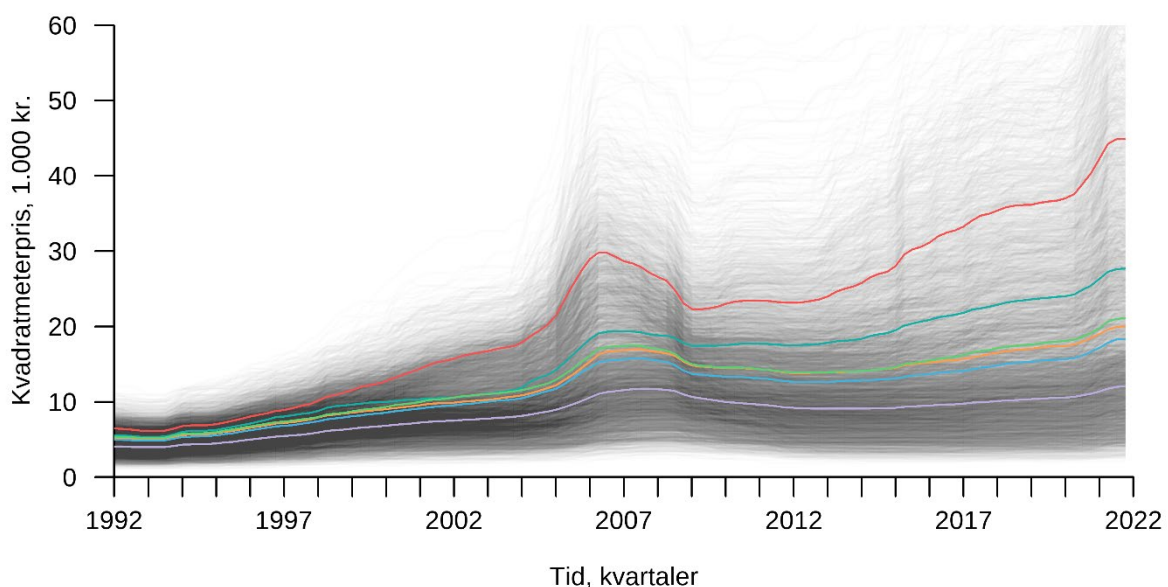
Figur 3.2 viser udviklingen i kvadratmeterprisen for 5.000 tilfældigt udvalgte ejerboliger. Det ses, hvordan der skønnes en individuel prisudvikling for hver bolig. Derudover er der indtegnet den gennemsnitlige kvadratmeterpris for forskellige bystørrelser.

På figuren er én bolig tegnet så svagt, at den knap fremgår. De mere tydelige grå linjer er således udtryk for, at prisskønnede for et antal boliger klumper sig sammen i grupper. Dette følger af, at prisudviklingen for nært beliggende ejerboliger vurderes næsten ens.

I begyndelse af 1990'erne ses en svagt faldende tendens i ejerboligprisen som blandt andet følger af kartoffelkuren og stigende arbejdsløshed. Herefter følger en periode med jævnt stigende boligpriser frem mod midt 2000'erne. Over få år følger en boligprisbølge med kraftige prisstigninger frem til omkring 2007. Herefter vendte udviklingen, og ejerboligprisen falder. Siden start 2010'erne har boligmarkedet igen været i fremgang. Særligt i hovedstadsområdet stiger boligprisen, mens stigningen i de mindre byer og landdistrikterne er mere afdæmpet.

Figur 3.2

Skønnet kvadratmeterpris for 5.000 tilfældigt udvalgte ejerboliger



Hovedstadsområdet Mindst 100.000 indbyggere 40.000-99.999 20.000-39.999 10.000-19.999 Færre end 10.000

Anm.: I figuren viser hver af de grå linjer den vurderede udvikling i kvadratmeterprisen for en tilfældigt udvalgt ejerbolig. De farvede linjer viser den gennemsnitlige kvadratmeterpris i en by af den givne størrelse.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

## 3.2 Regional udvikling i boligpriserne

I det følgende betragtes det, hvordan prisen for én boligkvadratmeter har udviklet sig siden 1992. Udviklingen i medianprisen betragtes, jf. Boks 3.2. Medianen er i mindre omfang end gennemsnittet påvirket af, at relativt få boliger gennem perioden opnår en meget høj kvadratmeterpris. Fordelingen af kvadratmeterprisen gennem perioden beskrives i afsnit 3.3.

Siden 1993 har boligpriserne haft en generelt stigende tendens. På landsplan er prisen i dag cirka 4 gange større end i begyndelsen af 1990'erne. Prisstigningen har været endnu større for ejerlejligheder. Dette har bidraget til, at ejerlejligheder i dag typisk er dyrere end parcelhuse pr. boligkvadratmeter. Især region Hovedstaden har oplevet store prisstigninger, jf. Figur 3.4.

Undervejs i perioden ses kraftige udsving i boligpriserne. I perioden 1992-2004 var ejerboligprisen støt stigende. Herefter skete en voldsom prisstigning i løbet af ganske få år midt i



00'erne. Det der skete i disse år, er efterfølgende blevet kategoriseret som en periode, hvor prisstigningerne ikke kan forklares ud fra den økonomiske udvikling i samfundet (også kaldet en "boligboble").

Udviklingen vendte brat, og boligejerne oplevede store kapitaltab i 2007-09. Nedturen på boligmarkedet begyndte i Københavnsområdet i andet halvår af 2006. Det var før virkningerne af den finansielle krise for alvor ramte økonomien. Faldende boligpriser bredte sig til resten af landet, og frem til 2009 skete betydelige prislefald. Især boligejerne i hovedstadsområdet var i denne periode hårdt ramt af store kapitaltab.

### Boks 3.2

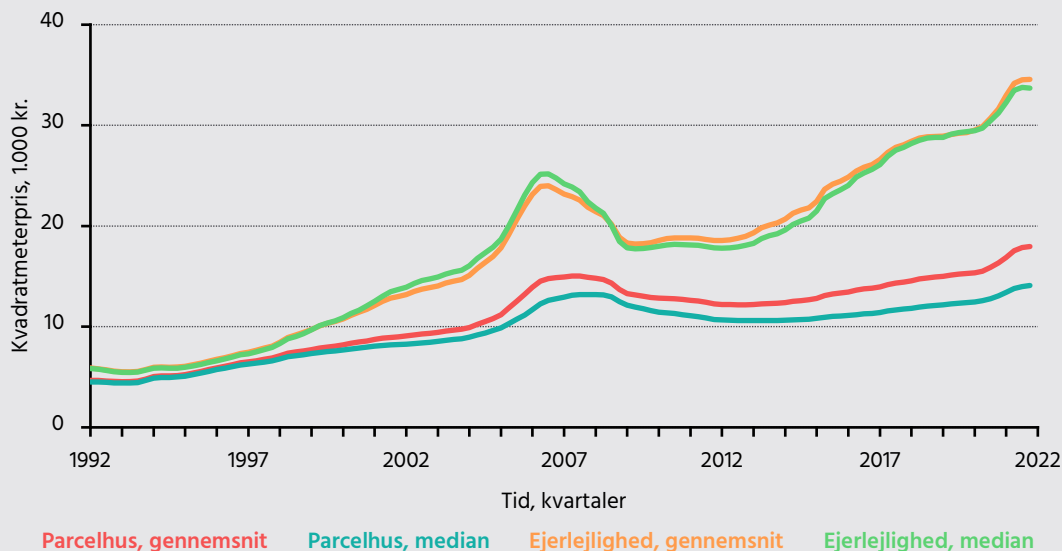
#### Kvadratmeterpris gennemsnit og median

Sædvanligvis anvendes den gennemsnitlige kvadratmeterpris til at illustrere boligprisudviklingen. Idet gennemsnittet påvirkes af relativt få boliger med meget høj kvadratmeterpris, kan det være at foretrække at betragte udviklingen i medianprisen i stedet. Medianen deler boligerne i to og angiver prisen på det midterste hus. Derfor er medianen mindre påvirket af få ekstreme værdier. En ulempe ved dette mål kan være, at medianboligens placering kan skifte over tid inden for det område, som betragtes.

På landsplan øges gennemsnitsprisen mere end medianprisen gennem den betragtede periode, jf. Figur 3.3. Dette afspejler, at der i løbet af perioden ses en større spredning i priserne, hvor forholdsvis få huse opnår en relativt høj pris. Tendensen er mere udtalt blandt parcelhuse end blandt ejerlejligheder, idet parcelhusene er mere uens i forhold til boligkarakteristika og placering.

Figur 3.3

#### Prisudvikling gennemsnit og median



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

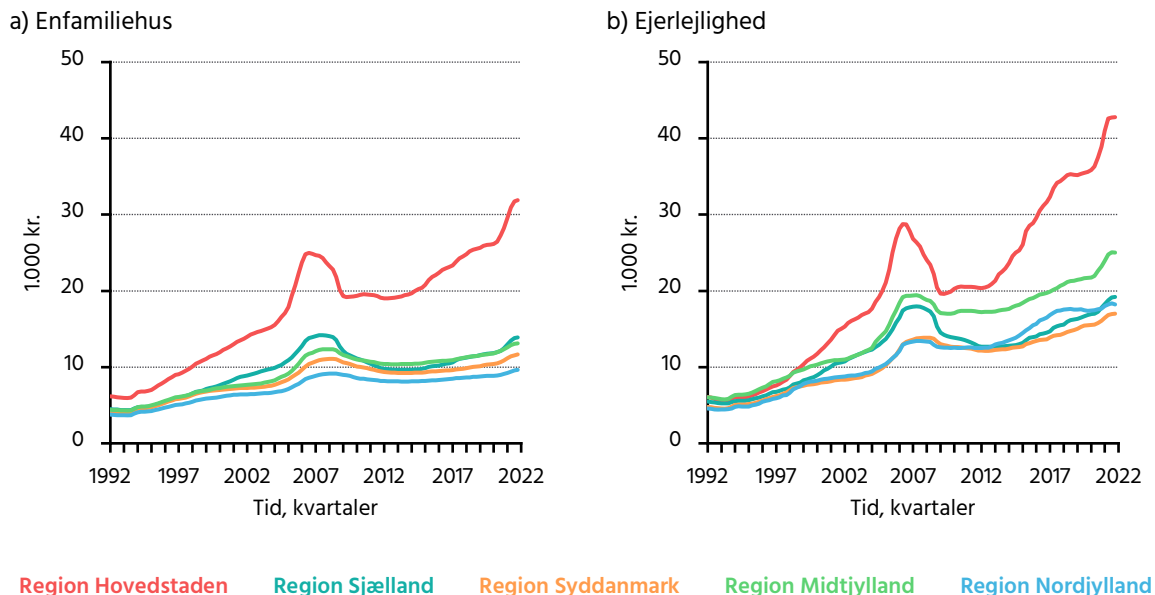
Herefter skete en stabilisering af boligprisen frem mod 2013, dog med store regionale forskelle.

De seneste ti år er ejerboligprisen igen steget ganske betragteligt. Prisstigningen har været særlig stor i hovedstadsområdet, men også i de største provinsbyer ses ejerboligprisen at være opadgående. Disse prisstigninger kan skyldes det ekstremt lave renteniveau, som sammen med en generel indkomstfremgang og en fortsat høj befolkningstilvækst presser særligt boligpriserne i hovedstadsområdet i opadgående retning.

Under coronapandemien accelererede boligprisstigningerne. Dette kan være forårsaget af en øget købekraft på grund af lave renter og flyttet efterspørgsel fra eksempelvis udlandsrejser til boligforbrug. Desuden kan pandemien have påvirket befolkningens præferencer for boligforhold og have fremrykket bolighandler i tid.

De stigende priser siden 2013 bredte sig fra de større byer til også at omfatte mindre byområder og landdistrikterne. Dog ikke i samme omfang som observeret frem mod finanskrisen.

**Figur 3.4**  
Kvadratmeterpris efter region



Anm.: Figuren viser den skønnede, gennemsnitlige handelspris pr. kvadratmeter for alle ejerboliger i regionen for hvert kvartal i perioden fra første kvartal 1992 til fjerde kvartal 2021.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

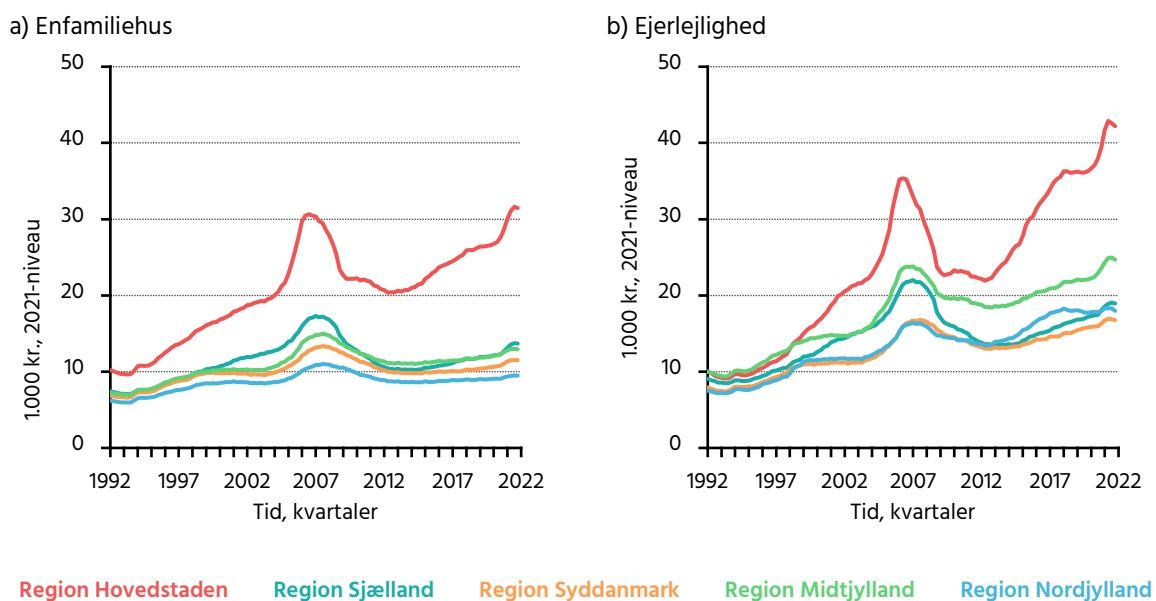
Udviklingen på ejerboligmarkedet siden begyndelsen af 1990'erne kan karakteriseres ved to overordnede tendenser.

For det første er boligpriserne siden 1992 steget mere end den almindelige inflation, jf. Figur 3.5. En række forhold kan være medforklarende til, at boligpriserne steg hurtigere end den normale prisstigningstakt. Heriblandt blev det fra første halvdel af 1990'erne muligt at finansiere boligkøbet med 30-årige lån mod hidtil 20 år, ligesom det igen blev muligt at optage rene annuitetslån. Selvom muligheden for rentetilpasningslån blev givet i 1996, blev låntypen for alvor udbredt efter årtusindeskiftet. Gennem 00'erne blev ligeledes givet mulighed for nye finansieringsformer. I 2003 gives muligheden for afdragsfrie lån, der hurtigt vinder

stor udbredelse blandt boligejerne. Muligheden for billigere finansiering samt et faldende renteniveau siden først i 1990'erne kan have været med til at presse boligpriserne op.

I 2002 laves desuden skattestop på ejendomsværdiskatten, så skattebetalingen siden har været fastfrosset på niveauet fra begyndelsen af årtusindeskiftet. At skattebetalingen ikke fulgte med op, da boligpriserne steg kraftigt midt i 00'erne og siden midt 2010'erne, kan have accelereret prisstigningerne yderligere. Modsat ville skattebetalingen falde, når boligpriserne går ned som det for eksempel var tilfældet i 2007-08, hvilket kunne have virket dæmpende på prisfaldet. I fravær af skattestoppet ville ejendomsværdiskatten altså have fungeret som en stabiliserende boligskat, der ville have dæmpet udsvingene på boligmarkedet.

**Figur 3.5**  
Real kvadratmeterpris efter region



Anm.: Figuren viser boligprisudviklingen korrigeret for inflation ved anvendelse af Danmarks Statistiks forbrugerprisindeks.  
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

For det andet er ejerboligprisen i Københavnsområdet steget kraftigere, end det var tilfældet i resten af landet. Denne tendens illustreres i Figur 3.6a, hvor udviklingen i ejerboligprisen i København sammenlignes med prisudviklingen omkring København og med kommunerne indeholdende Danmarks fem største byer foruden København.

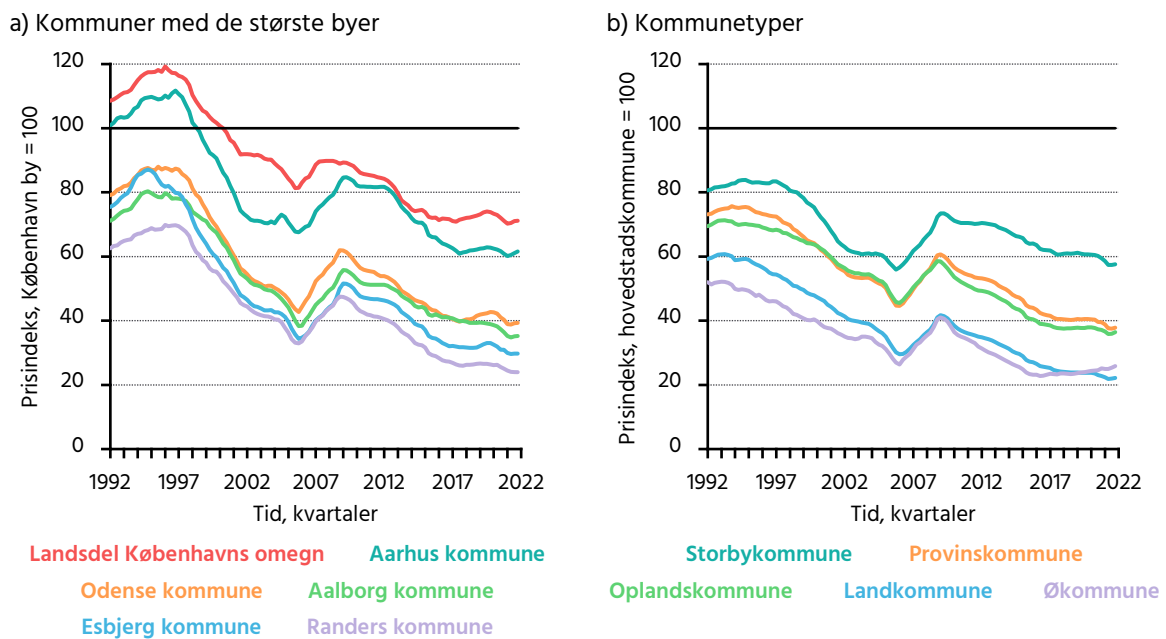
I 1990'erne var en ejerbolig i Københavns omegn eller i Aarhus op mod 20 pct. dyrere end en ejerbolig i landsdelen byen København (bestående af de fire kommuner København, Frederiksberg, Dragør og Tårnby). At ejerboligprisen i sidste halvdel af 90'erne steg relativt mere i København, end i de øvrige dele af landet, gjorde imidlertid, at prisniveauet i Københavns omkring årtusindeskiftet var på niveau med omegnskommunerne og Aarhus.

Gennem det nuværende årtusinde ses en tendens til, at ejerboligpriserne i København accelererede yderligere fra de andre største byområder i landet. Dog med tendens til nogenlunde samme stigningstakt i København og de øvrige storbyer siden 2017. I dag vurderes ejerboligprisen at være godt 20 pct. lavere i Københavns omegn end i byens centrum. Kvadratmeterprisen i Aarhus vurderes til omkring to tredjedele af prisen i København. I Odense og Aalborg kommune er prisniveauet i dag omkring 40-45 pct. af det centrale København.

At prisstigningen er særlig stor i København har sandsynligvis mange forklaringer (beskrives nærmere senere i kapitlet). En blandt flere årsager kan være, at ejerboliger i København udgør en forholdsvis lav andel af den samlede boligbestand. Ved stor efterspørgsel efter boliger i Københavnsområdet kan dette, i kombination med begrænset udbud af nye byggegrunde, betyde, at ejerboligprisen presses yderligere op end i øvrige egne af landet, da kun en mindre del af det samlede antal boliger kan omsættes frit på markedet. Dette kan være medforklarende til den observerede udvikling.

Hver kommune i landet typebestemmes ud fra antallet af indbyggere i kommunens største by og tilgængelighed til job, jf. Boks 3.3. Tendensen til, at ejerboligprisen i hovedstadskommunerne er blevet relativt dyrere end i øvrige dele af landet genfindes i Figur 3.6b, hvor grupperingen af kommunerne anvendes. Prisforskellen i forhold til hovedstadskommunerne er særligt øget i land- og økommunerne.

**Figur 3.6**  
Forskel i kvadratmeterpris i forhold til København



Anm.: Grupperingen af kommuner i figuren til højre er beskrevet i Boks 3.3.  
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Ejerboligprisen er ikke kun øget i hovedstadsområdet i forhold til øvrige dele af landet. Særligt efter år 2010 ses samme tendens ligeledes for de øvrige regioner. Her har boligprisen i regionens bymæssige centrum trukket fra boligprisen i omkringliggende områder. Det fremgår af Figur 3.7, hvor boligprisen i hovedstadsområdet, Odense, Esbjerg/Trekantsområdet, Aarhus og Aalborg holdes op mod de omkringliggende kommuner.

En række faktorer kan være medforklarende til, at ejerboligprisen i København og de øvrige storbyer i løbet af de seneste 30 år trak fra de øvrige dele af landet. I perioden ses en overordnet demografisk tendens til tilflytning mod de større byområder og deraf følgende reduktion i efterspørgslen efter boliger i de såkaldte yderområder. Dette presser prisen på ejerboliger i opadgående retning i byområderne.

### Boks 3.3

#### Anvendte områdeinddelinger

---

I rapporten anvendes fem forskellige geografiske opdelinger af landet i mindre enheder:

- De to administrative enheder regioner og kommuner,
- En statistisk opdeling i 11 landsdele, som er en underopdeling af de fem regioner,
- En opdeling efter størrelse af det byområde, hvor en given bolig er beliggende,
- En typebestemmelse af hver kommune som økommune, landkommune, oplandskommune, provinskommune, storbykommune eller hovedstadskommune ud fra antallet af indbyggere i kommunens største by og tilgængelighed til jobs.

En by defineres som en naturligt sammenhængende bebyggelse med mindst 200 indbyggere. Med naturligt sammenhængende menes, at afstanden mellem husene normalt ikke overstiger 200 meter med mindre afbrydelsen skyldes større gennemgående veje, sportspladser, grønne områder og lignende. Ved bystørrelse opgøres hovedstadsområdet samlet, mens øvrige byer grupperes efter antal indbyggere.

De anvendte områdeinddelinger beskrives nærmere i appendiks C.2.

Samtidigt har positiv nettoindvandring, hvoraf flertallet bosætter sig i de større byområder, øget boligefterspørgslen. Særligt i Københavnsområdet har positiv nettoindvandring fra udlandet medført øget pres på boligmarkedet. Et positivt fødselsoverskud har ligeledes været med til at øge befolkningstilvæksten.

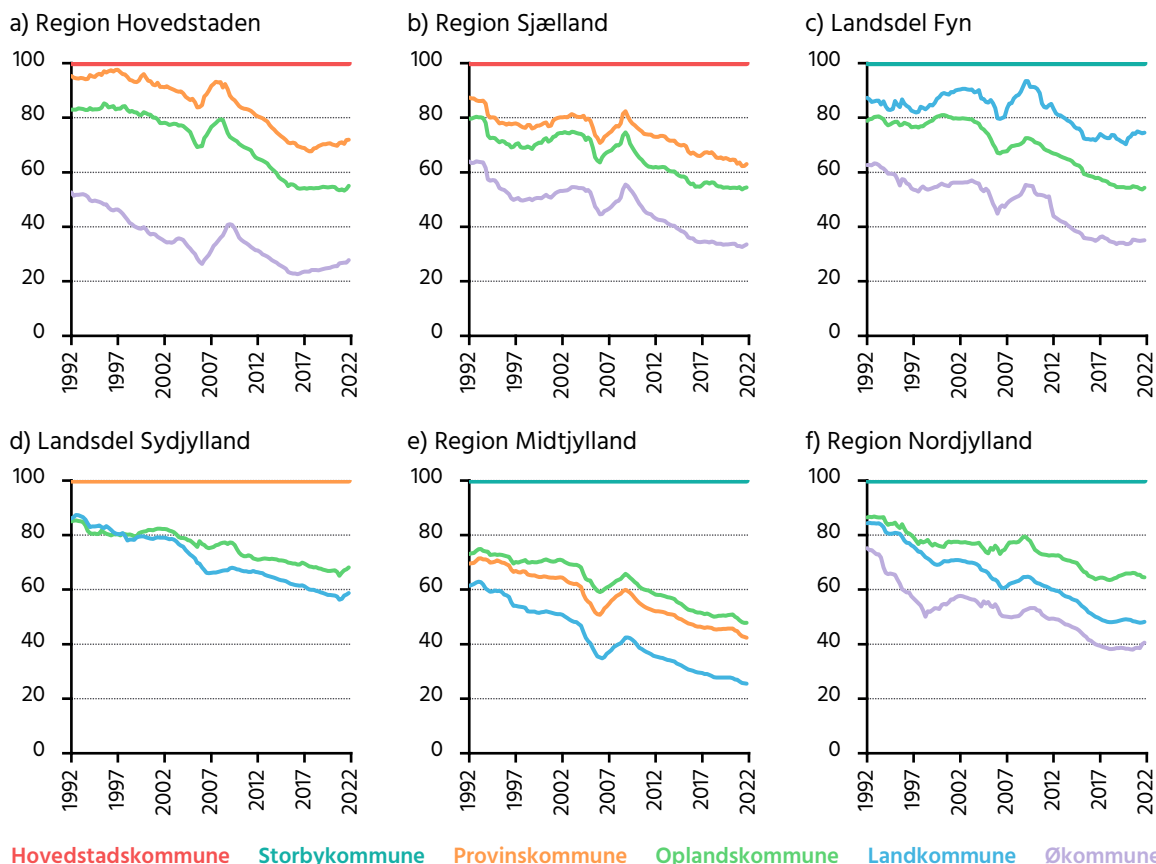
Andelen af unge, som fuldfører en videregående uddannelse, er desuden relativt høj det seneste årti. Da job, som efterspørger højt kvalificeret arbejdskraft, typisk er beliggende i de større byer, kan dette sammen med ændrede erhvervsstrukturer, hvor traditionelle erhverv forsvinder, have været med til at øge presset på ejerboligprisen i de større byer.

I de større byer er desuden begrænsede muligheder for at øge udbuddet af både ejerboliger og øvrige boligtyper, idet der typisk ses mangel på byggegrunde tæt ved centrum.

At boligpriserne især i hovedstæderne trækker fra det øvrige land observeres ikke udelukkende i Danmark. Samme tendens ses for eksempel også i de øvrige nordiske lande. I Norge og Sverige har effekten endda været stærkere, end det hidtil er set i Danmark.

Figur 3.7

Forskel i kvadratmeterpris i forhold til største by i området



Anm.: Førsteaksen viser kvartaler for perioden 1992-2021. På andenaksen vises et prisindeks, hvor den gennemsnitlige pris i den givne kommunetype sættes i forhold til de største bykommuner i det betragtede område. I figur a) og b) er referencegruppen lig hovedstadskommuner i den betragtede region, mens det i de øvrige figurer er Odense kommune (figur c), Esbjerg, Fredericia, Kolding og Vejle (figur d), Aarhus (figur e) samt Aalborg (figur f). Inddelingen af kommuner kan ses i appendiks C.2.

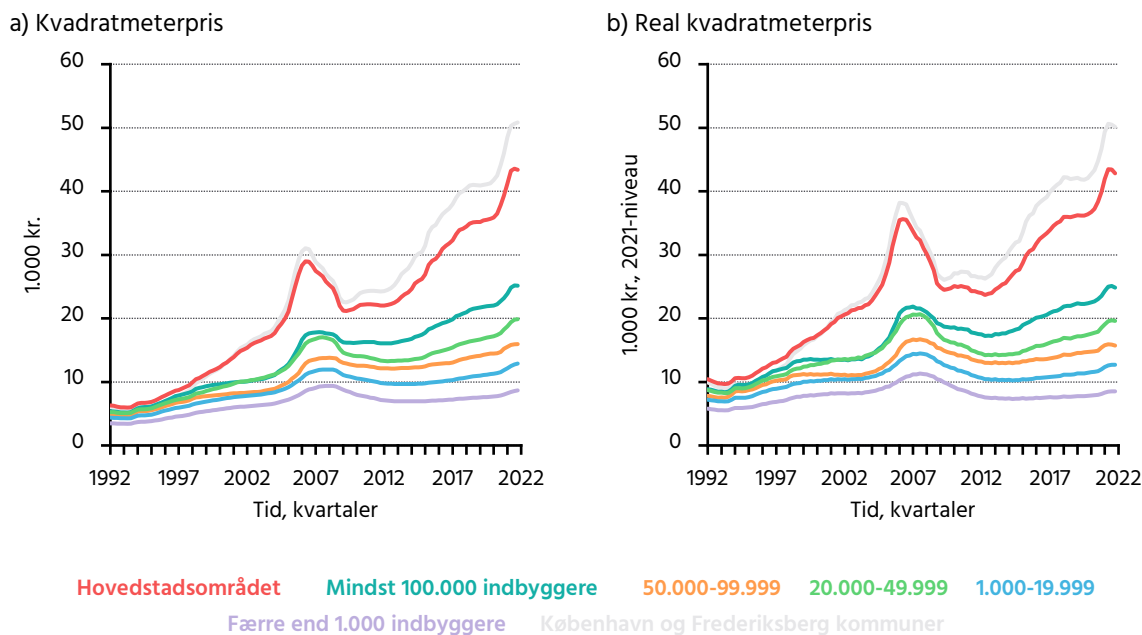
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Der ses en klar rangering, hvis boligprisudviklingen opgøres efter bystørrelse, jf. Figur 3.8. Dels er boligprisen generelt højere, jo større byområde, som betragtes. Dels ses over perioden en kraftigere prisstigning i de større byområder end i de mindre, hvorfor forskellen mellem boligprisen i by- og landområder er øget betydeligt, jf. Figur 3.9. I de mindre landsbyer og i landdistrikterne er boligprisstigningen over det seneste årti beskeden, men der er observeret en kraftig stigning i de større byområder. Dette forklarer den øgede forskel mellem by og land.

Rangeringen efter bystørrelse, så medianprisen er højere jo større byområde der betragtes, er udbredt. Undtaget er dog region Midtjylland, jf. Figur 3.9. I denne region er der ikke væsentlig forskel på niveauet for boligprisen i byer med 50.000-99.999 indbyggere (Randers, Horsens og Herning), og byer med 10.000-49.999 indbyggere i form af Silkeborg, Skanderborg og Ikast. Prisniveauet i sidstnævnte gruppe trækkes op af relativt høje kvadratmeterpriser i Silkeborg og Skanderborg, hvor ejerboligprisen blandt andet er præget af nærheden til Aarhus.

Figur 3.8

Kvadratmeterpris efter bystørrelse



Anm.: Inddelingen efter bystørrelse er nærmere beskrevet i appendiks C.2. Den reale prisudvikling viser boligprisudviklingen korrigeret for inflation.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Den gennemsnitlige stigning i boligpriserne over de seneste 30 år dækker således over markante forskelle på tværs af landet. Den geografiske variation er forsøgt illustreret i Figur 3.10, hvor prisudviklingen på ejerboliger i udvalgte kommuner er afbildet. I nederste del af figuren vises ændringen i boligprisen i forhold til samme kvartal ét år tidligere.

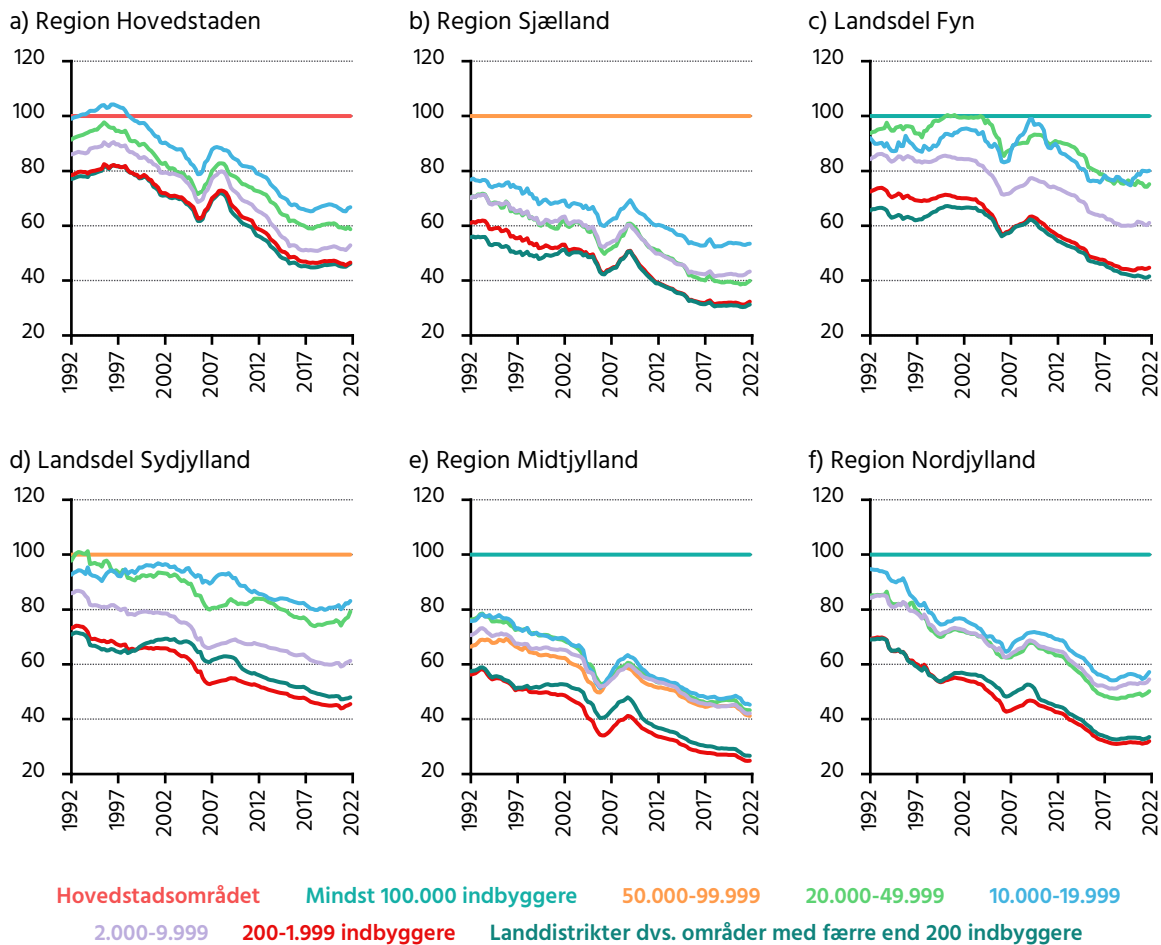
Fremgangen på boligmarkedet gennem første halvdel af 00'erne startede i Københavnsområdet, og prisstigningerne var størst i hovedstadsområdet og omkring Aarhus. Kommunerne nord for København, hvor parcelhusene er relativt dyre, oplevede kraftige nominelle prisstigninger. De stigende priser breder sig til de større provinsbyer, som ligeledes oplever pæne prisstigninger frem mod finanskrisen. I de mindre byer samt landområderne indtræffer udviklingen senere i tid og prisstigningen er her mere afdæmpet.

Nedgangen på boligmarkedet startede i andet halvår 2006 i landets to største byer. Herefter følger tre år frem mod andet kvartal 2009, hvor priserne i København og Aarhus falder voldsomt. Kvadratmeterprisen på ejerlejligheder i København falder 35 pct. (Aarhus 22 pct.) i denne periode. De faldende boligpriser breder sig til resten af landet. På landsplan toppes boligpriserne i første kvartal 2007, hvorefter bunden nås midt i 2009, hvor den gennemsnitlige kvadratmeterpris for hele landet er faldet med 19 pct. Generelt skete de største fald i boligpriserne i de områder af landet, der under opturen havde oplevet de største prisstigninger.

De seneste godt ti år har udviklingen i boligpriserne været meget forskellig i forskellige dele af landet. I Københavnsområdet, Aarhus samt en række større byer såsom Odense, Silkeborg, Skanderborg og Aalborg er boligpriserne gået frem. Fremgangen i disse områder startede forholdsvis tidligt efter finanskrisen. Særligt i perioden frem mod 2016 ses større prisstigninger i de største byområder i forhold til de mindre provinsbyer og landdistrikterne. Herefter ses prisstigningerne at følge i samme takt.

Figur 3.9

Forskel i kvadratmeterpris i forhold til største by i området efter bystørrelse



Anm.: Førsteaksen viser kvartaler for perioden 1992-2021. På andenaksen vises et prisindeks, hvor den gennemsnitlige pris i den givne bystørrelse sættes i forhold til den største bystørrelse i det betragtede område. I figur a er referencegruppen lig hovedstadsområdet, mens det i de øvrige figurer er byer med mindst 100.000 indbyggere (Odense i figur c, Aarhus i figur e og Aalborg i figur f), eller byer med 50.000-99.999 indbyggere (Roskilde i figur b og Esbjerg/Kolding/Vejle i figur d). Inddelingen efter bystørrelse er nærmere beskrevet i appendiks C.2.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

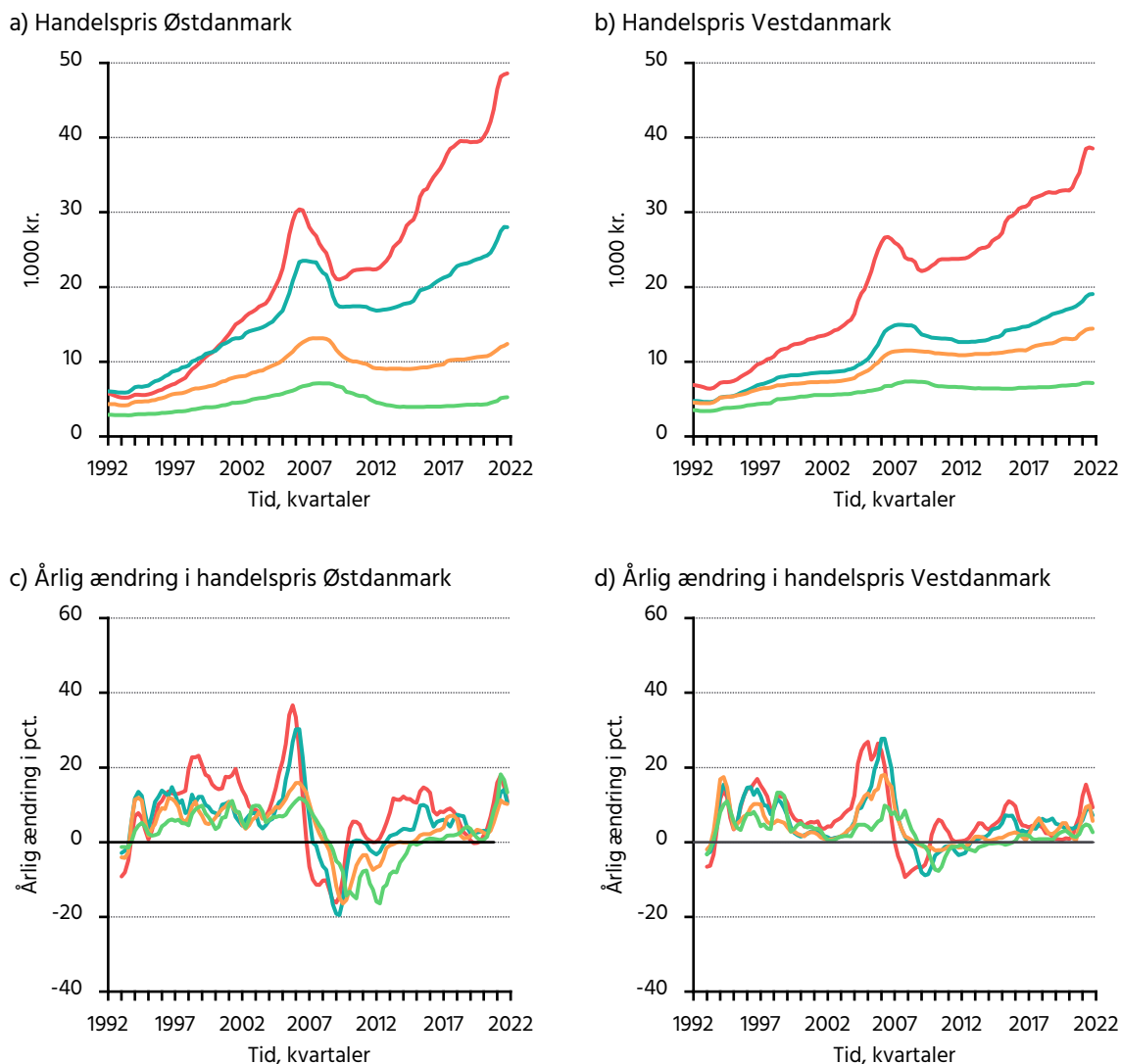
I perioden fra 2012 til omkring 2016 følger boligpriserne i en række kommuner som for eksempel Roskilde, Køge, Herning og Hjørring omtrent den almindelige prisudvikling. Andre kommuner oplever forholdsvis konstante priser, hvilket omfatter områder som Lejre, Stevns, Sønderborg, Fredericia, Randers samt Viborg kommune. I landkommunerne var der fra 2009 en årrække med stiltand på boligmarkedet. Dette gælder for eksempel Vest- og Sydsjælland, langs dele af Jyllands vestkyst og i Nordjylland.

I løbet af 2016-17 ses tegn på, at prisstigningerne breder sig fra de større byer til også at omfatte mindre byområder og landdistrikterne. I perioden herefter ses stigende priser i størstedelen af landet, og prisstigningerne accelerere under coronapandemien. I løbet af 2021 ses tendens til en stabilisering af priserne.



Figur 3.10

Kvadratmeterpris udvalgte kommuner



København, ejerlejlighed Roskilde Slagelse Lolland Aarhus, lejlighed Odense Esbjerg Vesthimmerlands

Anm.: Figuren viser udviklingen i den skønnede handelspris for alle ejerboliger i de udvalgte kommuner, i København og Aarhus kommuner dog kun ejerlejligheder. Den årlige ændring i handelsprisen er beregnet som prisen i det angivne kvartal sat i forhold til prisen fire kvartaler tidligere.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

### 3.3 Variation i boligprisen

Ved at betragte medianen for kvadratmeterprisen har undersøgelsen hidtil fokuseret på det centrale boligpriseniveau. For at belyse variationen omkring medianen fokuseres der i dette afsnit på fordelingen af kvadratmeterpriser også kaldet tæthedsfunktionen eller blot tæthed.

Tætheden er mål for, hvor koncentreret eller spredt ejerboliger er fordelt over kvadratmeterpriser, og i afsnittet bruges tæthedsplot til at identificere fordelingsstoppe, spredning og asymmetri. Sammenligning af fordelingerne over tid og i forskellige geografiske områder gør det muligt at identificere om usædvanlige mønstre i kvadratmeterpriserne findes.

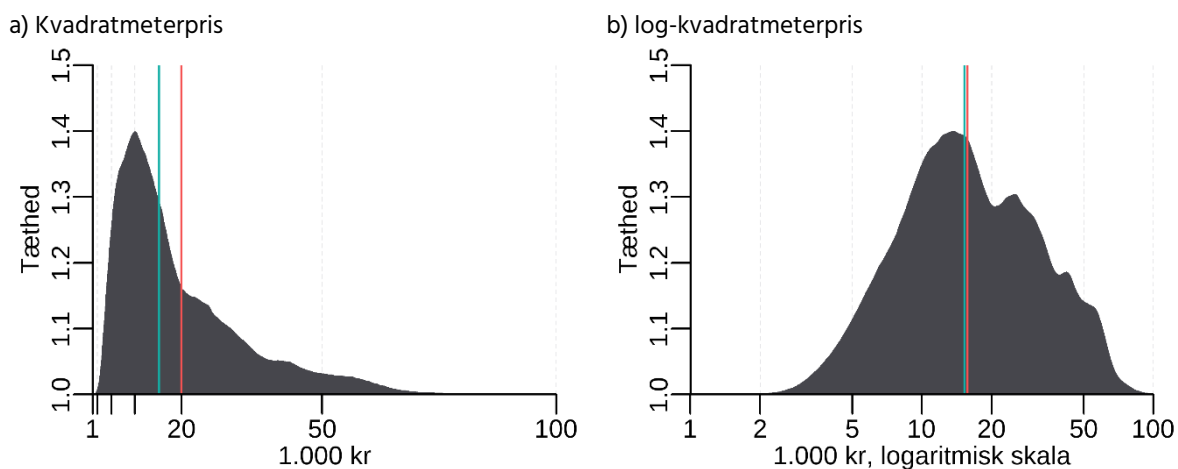
Figur 3.11 viser tætheden af ejerboliger over kvadratmeterpriser i 2021. Tætheden vises både med og uden logaritmisk transformation. Tæthed beskriver sandsynligheden for, at en ejerbolig har en bestemt kvadratmeterpris. Jo flere ejerboliger der har en pris i et givet interval, desto højere er tætheden i intervallet.

Figuren viser, at ejerboligerne er skævt fordelt over kvadratmeterpriser, og at gennemsnit, median og modus<sup>5</sup> ikke er sammenfaldende. Medianprisen er lavere end gennemsnitsprisen, hvilket betyder, at et større antal ejerboliger har en lavere kvadratmeterpris end gennemsnitsprisen. Gennemsnitsprisen bliver trukket op af relativt få ejerboliger med høje kvadratmeterpriser. Herved er medianen at foretrække som et mål for den centrale boligpris i denne analyse. Medianen deler ejerboligerne i to lige store grupper og er herved mindre påvirket af få ekstreme værdier end gennemsnittet.

Fordelingen af kvadratmeterpriser er mindre skævt fordelt på en logaritmisk skala og visuelt tættere på en normalfordeling<sup>6</sup>. Fordelen ved den logaritmiske transformation er, at den komprimerer større værdier og udvider mindre værdier. Datamønstre i den skæve fordeling bliver herved mere synlige. Den logaritmiske transformation er en monoton transformation, hvilket betyder, at den interne rækkefølge bevares. Desuden vil mediankvadratmeterprisen ikke ændres af transformationen. Resten af afsnittet omhandler logaritmisk transformeret kvadratmeterpris.

Figur 3.11

Fordeling af kvadratmeterpris, 2021



Anm.: Gennemsnit og median er markeret med henholdsvis rød og petroleumsfarvet lodret streg. Median kvadratmeterprisen er i begge figurer 13.918 kr., mens den gennemsnitlige kvadratmeterpris er 17.733 kr. Arealet under kurven er normaliseret til én, hvilket betyder, at det samlede areal under kurven repræsenterer det samlede antal ejerboliger.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur 3.12 er et såkaldt bønneplot (eng: beanplot, jf. Kampstra (2008)) af reale kvadratmeterpriser i perioden 1992-2021. Figuren viser den estimerede tæthedsfunktion for ejerboliger i Danmark ultimo 2021, og alle boliger optræder i hele perioden. Hver 'bønne' i figuren viser et ulige årstal i den venstre del og et lige år i den højre. Gennemsnit, median og kvartiler er

<sup>5</sup> Modus er toppunktet af fordelingen og angiver den kvadratmeterpris som flest ejerboliger har.

<sup>6</sup> Kolmogorov Smirnov test for (log) normalitet afvises. Testen sammenligner den empiriske fordeling af stikprøven med den teoretiske log normalfordeling

markeret med henholdsvis rød, petroleumsfarve og orange. Figuren giver mulighed for at sammenligne fordelingerne over tid. Herved kan det undersøges, om fordelings funktionelle form har været konstant, eller om den har ændret sig.

Fra figuren fremgår det, at fordelingerne ændres over perioden. Det betyder, at de tidligere beskrevne ændringer i medianpriser ikke blot er en parallelforskydning af fordelingerne.

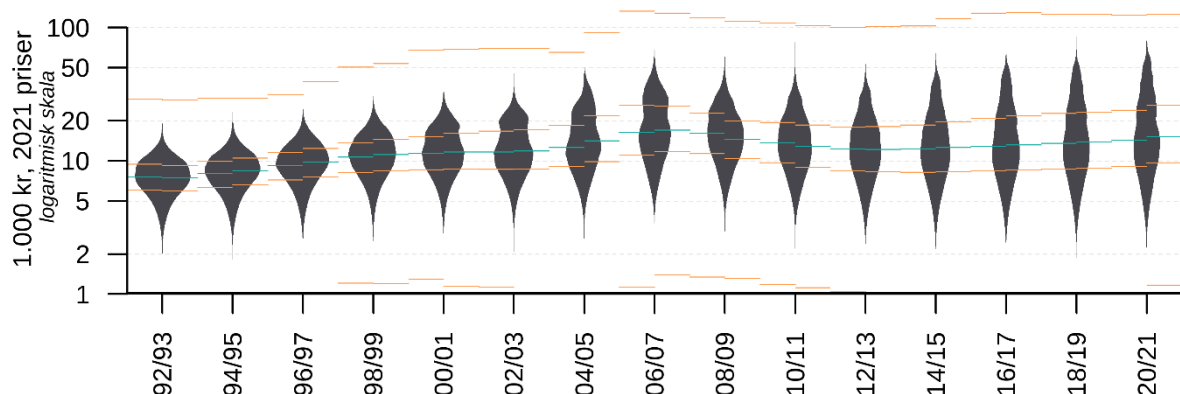
Det ses, at fordelingerne bliver bredere og mere udstrakte over tid. Mod slutningen af perioden ses det, at flere ejerboliger har kvadratmeterpriser, der afviger mere fra medianprisen, end det er tilfældet i starten af perioden. Over perioden oplever boligerne en stigende prisdiversitet. Det betyder, at variansen og spredning for kvadratmeterpriserne er vokset gennem perioden. Jo større spredningen er, desto større usikkerhed er der omkring median- eller gennemsnitsprisen.

Variationen i kvadratmeterpriserne følger i nogen grad konjunkturudviklingen. Spredningen på boligpriserne stiger op til finanskrisen i 2007 og falder i årene efter. Hvorefter den frem til 2021 igen er stigende. Dette indikerer at boligpriserne glider væk fra hinanden, og at ejerboligmarkedet bliver mindre homogent med hensyn til kvadratmeterpriser, i perioder med højkonjunktur.

Området mellem 1. og 3. kvartil (eng: interquartile range, IQR), markeret med orange i Figur 3.12, indeholder halvdelen af ejerboligerne svarende til de mest gennemsnitlige boliger. Det ses, at dette interval er blevet større over tid, hvorved der også er større variation i priserne for de mest gennemsnitlige boliger.

I perioden op til finanskrisen, rundt regnet perioden 2001 til 2007, ses en tendens til, at fordelingen bliver bimodal, dvs. får flere toppe eller flere modus. Modus er den kvadratmeterpris, der forekommer oftest. Fordelingstoppene er ikke lige høje, idet den kvadratmeterpris hvor toppen er højest, er den oftest forekommende. Den bimodale fordeling kan indikere, at der i perioden sker en opdeling i boligmarkedet i to eller flere delmarkeder, hvor hvert delmarked har forskellige median og gennemsnitspriser. Delmarkederne kunne være defineret af geografiske områder og bystørrelser.

**Figur 3.12**  
Fordeling af real kvadratmeterpris, 1992-2021



Anm.: Figuren viser den estimeret tæthedsfunktion for logaritmisk transformerede kvadratmeterpriser. Median og kvartiler er markeret med henholdsvis petroleumsfarve og orange. Fordelingerne til venstre og højre i hver 'bønne' er henholdsvis lige og ulige år.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

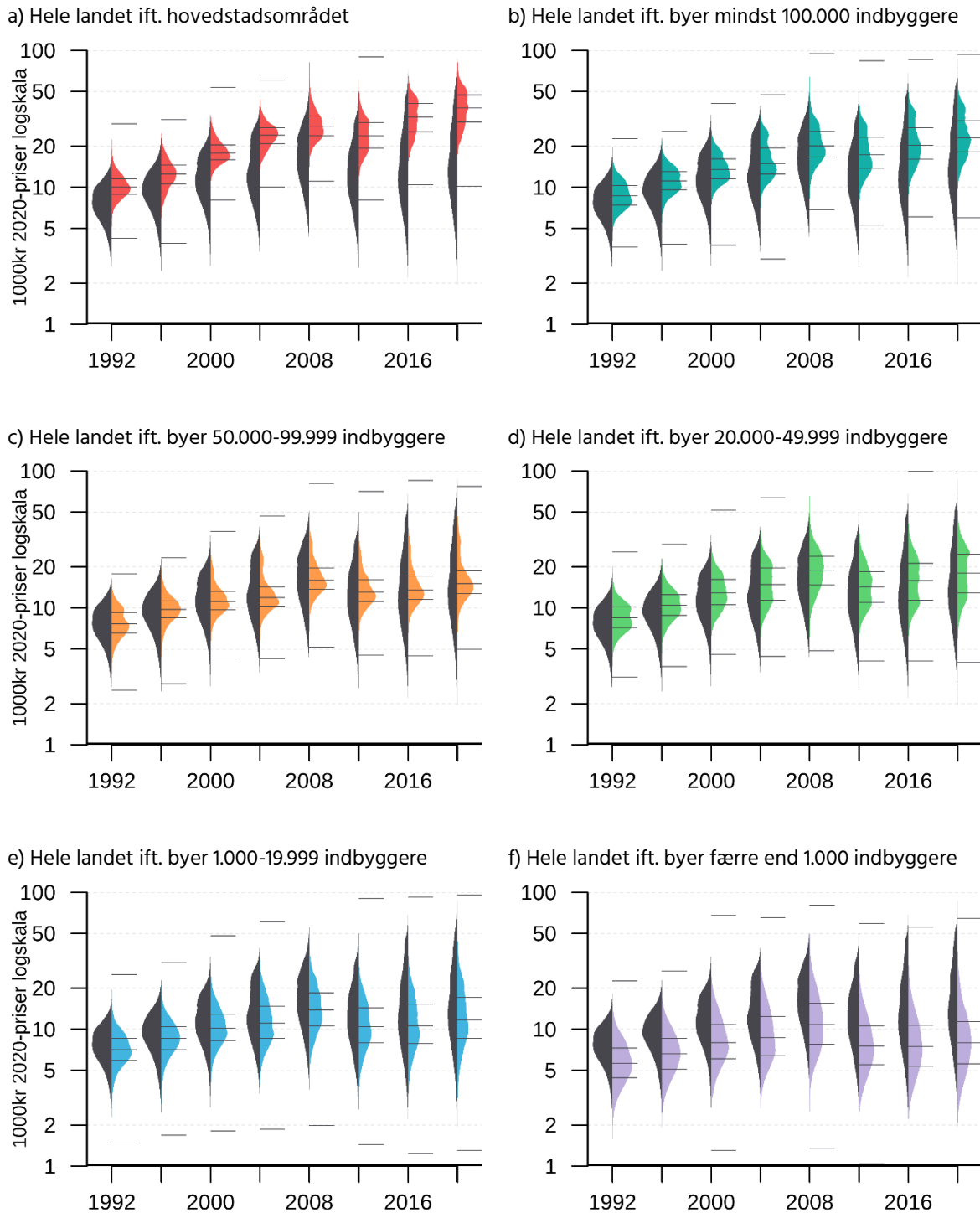
Mediankvadratmeterprisen for hele ejerboligmassen og modus er i perioden 2001 til 2007 relativt forskellige. Dette betyder, at mediankvadratmeterprisen i denne periode kun vil være repræsentativ for relativt få boliger og derfor skal anvendes som en central tendens med forsigtighed.

Opdeles ejerboligmarkedet efter bystørrelse, kan eventuelle forskelle i markedstendenser identificeres. Figur 3.13 viser tætheden for logkvadratmeterpriser efter bystørrelse og for hele landet. Generelt ses samme tendenser for de forskellige bystørrelser som for hele landet: Fordelingerne bliver bredere og mere udstrakte over tid, variansen i logkvadratmeterprisen er voksende for alle bystørrelser.

For den enkelte bystørrelse gælder, at boligmasse er mere homogen med hensyn til logkvadratmeterprisen end Danmark som helhed. Variationen omkring medianen er mindre for store byer, end det er tilfældet for de mindre byområder. En del af forklaringen herpå er, at større byer ofte har en mere ensartet boligmasse med hensyn til boligtype, størrelse, geografisk placering, opførselsår og byggestil. Derudover kan der også være mere ensartet efterspørgsel fra købere i større byer. På den anden side har landdistrikter en mindre homogen boligmasse, da den dækker et større geografisk område. For eksempel kan priserne på boliger nær kysten have udviklet sig anderledes end priserne på ejerboliger centrale placeret i landet.

Figur 3.13

Fordeling af real kvadratmeterpriser efter bystørrelse, udvalgte år 1992-2021



Anm.: Hele landet afbildet som grå, mens den betragtede bystørrelse er farvet. Arealet under tæthedsfunktionen er lig 1, hvorfor tæthedsfunktionen for forskellige populationsstørrelser kan sammenlignes. Gennemsnit og kvantiler er markeret med henholdsvis mørk og lys grå streg.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik



## 4. Geografisk variation i ejerboligpriser

Der er estimeret en kvadratmeterpris for alle ejerboliger i perioden 1992-2021. Dette giver mulighed for at undersøge den geografiske variation i udviklingen i ejerboligprisen på et meget detaljeret niveau.

Over de seneste 30 år er boligprisen i København trukket fra de øvrige dele af landet. Samme tendens observeres omkring de øvrige regionale bycentre, jf. kapitel 3.

Den geografiske variation i boligprisudviklingen undersøges ved at betragte den skønnede prisudvikling for hver enkelt ejerbolig. Metoden til at lave disse skøn er opsummeret i afsnit 3.1 og grundigt beskrevet i kapitel 5.

Det undersøges hvilke geografiske områder, der i perioden siden 1992 har oplevet store prisstigninger og hvilke dele af landet, som ikke i samme grad har oplevet en opadgående pris. Herved illustreres det, i hvilke området af landet boligprisen trak fra andre områder.

Udviklingen illustreres på landkort<sup>7</sup>. Derfor betragtes boligprisen for en standard ejerbolig med den unikke geografiske beliggenhed som faktisk observeret.

### Boks 4.1

#### Standardbolig

Standardejerboligen er defineret ud fra oplysninger om den typiske ejendom. Der skelnes mellem huse og lejligheder.

Standardhuset er et parcelhus opført i 1970. Det har ydervæg af mursten, cementstentag, ét badeværelse, ét toilet, fjernvarme, ingen supplerende varme og boligareal på 135 m<sup>2</sup> plus 12 m<sup>2</sup> kælder. Huset er beliggende på en grund på 845 m<sup>2</sup> og har almen offentlig vandforsyning. Ejerne har beboet boligen i 14 år og er i 7. indkomstdecil.

Standardlejligheden er en ejerlejlighed opført i 1970. Lejligheden er beliggende i en murstens-ejendom med fladt tag opført i 1970, som har fjernvarme uden supplerende varmekilde og alment offentlig vandforsyningsanlæg. Lejligheden har ét bad og ét toilet. Ejerne har beboet lejligheden i syv år og er i 7. indkomstdecil.

Det er kun bygnings- og ejerkaraktistika, som fastsættes til referenceværdier. Geografiske karakteristika såsom kommune, bystørrelse og afstandsvariable til eksempelvis motorvej og kyst fastholdes på deres faktiske værdier, hvorfor disse vil variere mellem standardboligerne.

<sup>7</sup> For at leve op til gældende diskretionskrav er hele landet blevet inddelt i områder (polygoner) ved en ny metode, som er udviklet til formålet. Hvert område indeholder mindst 50 husstande og mindst 100 personer. Inddelingen foretages inden for hver kommune, så et område ikke overskrider kommunegrænser. Samtlige ejerboliger i et sådan område farvelægges ens, hvorved diskretionskravet efterleves.

Men standardbolig menes, at karakteristika for boligen og dens ejere fastsættes til enslydende værdier for alle boliger. Geografiske karakteristika fastholdes som unikke for hver bolig, idet det netop er betydningen af placeringen, der ønskes undersøgt. Standardboligen defineres i Boks 4.1.

Ved at anvende den samme standardiserede ejerbolig kan forskelle i ejendomsværdien pr. kvadratmeter henføres til forskelle i geografiske forhold.

## 4.1 Identifikation af dyre og billige områder

Ud fra de estimerede kvadratmeterpriser er det muligt at identificere den geografiske placering af forskellige prisområder. Undersøgelsens første år 1992 sammenlignes med seneste datoår 2021.

Prisområderne dannes ved at rangordne alle ejerboliger efter deres skønnede kvadratmeterpris. Herefter inddeles boligerne i ti lige store grupper (kaldet deciler). Alle boliger indgår i hele perioden (uanset opførelsesår). Det betyder, at der i alle år er samme antal boliger i hvert decil. De ti pct. ejerboliger med lavest kvadratmeterpris tilhører første decil og er vist i Figur 4.1. Gruppen med de dyreste boliger (tiende decil) vises i Figur 4.2. Tilsvarende figurer for alle ti deciler vises i appendiks A.

De ejerboliger med de ti pct. laveste kvadratmeterpriser er geografisk spredt ud over det meste af Danmark fraregnet Sjælland. Desuden ses et fravær af de billigste boliger i et område med centrum i Aarhus, som også dækker Silkeborg og Skanderborg. Dette samme gør sig gældende omkring de øvrige større byer herunder blandt andet Odense, Aalborg, Trekantsområdet, Esbjerg, Herning, Holstebro og Viborg.

Den geografiske koncentration af første decil og områders placering har ændret sig over tid. I oplandet til de større byer ses en tendens til, at de billigste boliger forsvinder gennem perioden. Tendensen er mest tydelig på Fyn, hvor boliger i første prisdecil er forsvundet fra en cirkel med centrum i Odense.

Det samme gør sig gældende omkring Aarhus, Trekantsområdet og Aalborg, dog i mindre udtalt grad end omkring Odense. Dette leder samlet set til, at der ses færre billige ejerboliger i området mellem Vejle, Horsens og Silkeborg, et område nord for Aarhus og syd Randers, et område sydøst for Aalborg og endelige området omkring Odense som i 2021 er flyttet længere væk fra Odense centrum.

Modsat opstår der i 2021 andre områder med ejerboliger med en relativt lav kvadratmeterpris. Disse områder opstår typisk i de tyndere befolkede områder og dækker primært et område omkring Grenå på Djursland, området omkring Hjørring i Nordjylland, et område syd for Ringkøbing fjord og endelige et område ved Bredebro i Sønderjylland.

Ejerboligerne med de højeste kvadratmeterpriser er geografisk sammenklynget omkring Aarhus, København og Nordsjælland. For sidstnævnte gælder det særligt boliger beliggende ved kysten. Det ses en tendens til, at boliger i tiende decil i løbet af perioden 1992-2021 bliver mere sammenklynget omkring centrum af de to største byer.

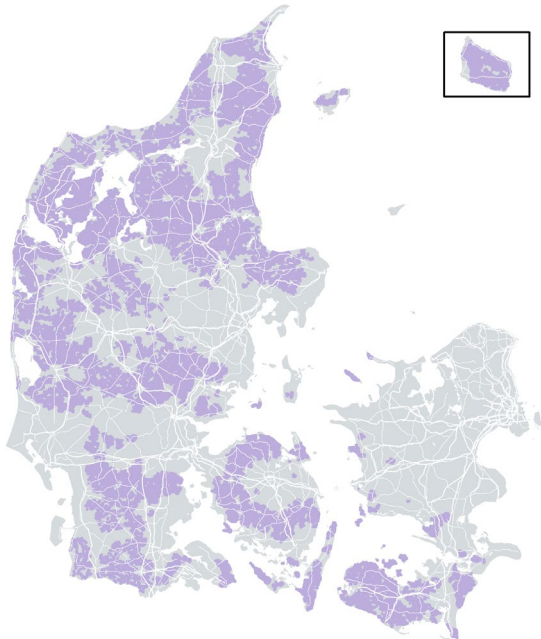
I det følgende delafsnit ses der nærmere på, hvor meget kvadratmeterprisen i de berørte områder har ændret sig og hvilken prisdeciler ejerboligerne nu tilhører eller kommer fra. Desuden betragtes udviklingen i den geografisk koncentration af ejerboligpriser.



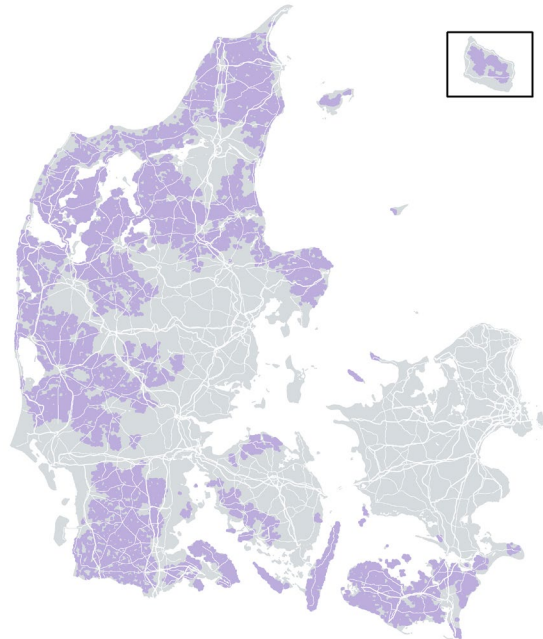
**Figur 4.1**

**Placering af de ti pct. ejerboliger med lavest kvadratmeterpris**

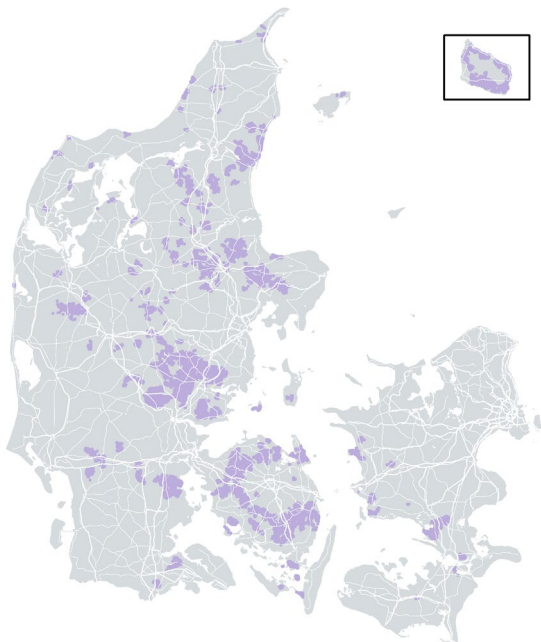
a) 1992



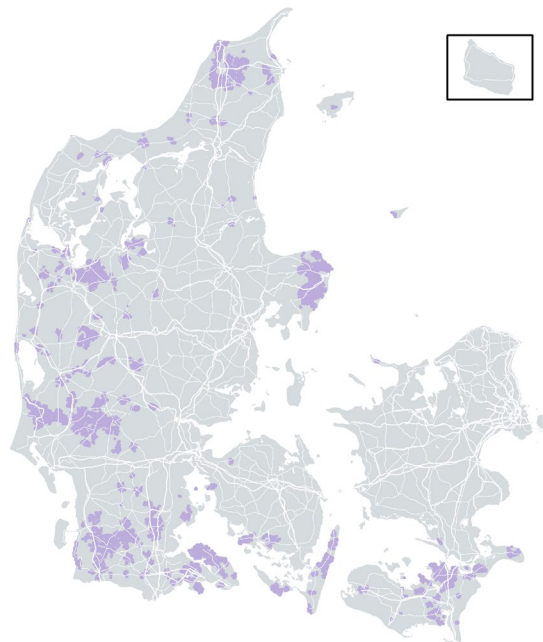
b) 2021



c) Forsvundne områder i 2021



d) Nye områder i 2021



Anm.: Alle ejerboliger er rangeret efter boligens skønnede kvadratmeterpris. Figuren viser placeringen af de ti pct. af boligerne med den laveste kvadratmeterpris. Desuden vises boliger som forlader og tilkommer gruppen i perioden 1992-2021.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

**Figur 4.2**

**Placering af de ti pct. ejerboliger med højest kvadratmeterpris**

a) 1992



b) 2021



c) Forsvundne områder i 2021



d) Nye områder i 2021



Anm.: Alle ejerboliger er rangeret efter boligens skønnede kvadratmeterpris. Figuren viser placeringen af de ti pct. af boligerne med den højeste kvadratmeterpris. Desuden vises boliger som forlader og tilkommer gruppen i perioden 1992-2021.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

### Prismobilitet

For alle ejerboliger viser Figur 4.3 prismobiliteten mellem 1992 og 2021. Figuren beskriver, hvor stor en andel af ejerboliger, der tilhører et bestemt decil i 1992, som stadig er i denne eller er skiftet til et anden decil i 2021.

Er matricen diagonalt centreret, indikerer det, at der har været begrænsede bevægelser mellem deciler. Hvis matricen er fuldstændig diagonalt centreret, betyder det, at alle ejendomme er i samme decil i 2021, som de var i 1992.

Generelt fremgår det, at ejerboligerne i 2021 placeres indenfor 30 pct. af det oprindelige decil i 1992. Der er en tendens til, at de billigste og dyreste ejerboliger har oplevet den mindste prismobilitet. 65 pct. af de billigste ejerboliger i 1992 er stadig blandt de ti pct. billigste ejerboliger i 2021. 89 pct. af dem er blandt de 20 pct. billigste ejerboliger.

I den modsatte ende af skalaen ses det, at 47 pct. og 79 pct. af de dyreste ejerboliger i 1992, stadig i 2021 er blandt henholdsvis de henholdsvis ti og 20 pct. dyreste ejerboliger. For de centralt rangerede ejerboliger ses en lidt større prismobilitet både op og ned i prisgrupper.

**Figur 4.3**

Mobilitet mellem prisdeciler, pct.

		2021									
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
1992	D1	65	24	8	2	1	0	0	0	0	0
	D2	25	33	23	12	5	2	1	0	0	0
	D3	8	22	26	20	13	7	3	0	0	0
	D4	2	11	20	23	19	14	7	2	1	1
	D5	1	5	12	18	21	19	15	5	2	3
	D6	0	2	6	11	17	19	21	11	5	5
	D7	0	1	3	7	12	16	20	21	11	9
	D8	0	0	1	4	8	13	16	25	20	14
	D9	0	0	1	1	4	9	14	22	29	21
	D10	0	0	0	0	1	2	4	13	32	47

Anm.: Alle ejerboliger er rangeret efter boligens skønnede kvadratmeterpris og inddeles i deciler. Første decil D1 indeholder boligerne med lavest kvadratmeterpris. Figuren viser prisdecil for boligen i 1992 (nedad) og i 2021 (udad).

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur 4.4 viser mobiliteten mellem deciler efter bystørrelse. Inddelingen i deciler opgøres fortsat på landsplan, hvorfor der inden for en bystørrelse ikke er lige mange boliger i hvert decil. Antal ejerboliger i hvert decil i 1992 angives i firkantede parenteser (angivet i tusind boliger).

Figur 4.4

Mobilitet mellem prisdeciler efter bystørrelse, pct.

a) Hovedstadsområdet

		2021										
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	
1992	D1	[0]	0	0	0	0	0	0	0	0	91	9
	D2	[0]	0	0	0	0	1	1	6	5	58	29
	D3	[1]	0	0	0	0	0	2	19	9	37	32
	D4	[3]	0	0	0	0	0	1	14	16	28	42
	D5	[8]	0	0	0	0	0	1	15	15	25	44
	D6	[17]	0	0	0	0	0	0	10	21	26	43
	D7	[29]	0	0	0	0	0	0	3	26	30	42
	D8	[39]	0	0	0	0	0	0	1	20	33	46
	D9	[45]	0	0	0	0	0	0	0	9	37	53
	D10	[65]	0	0	0	0	0	0	0	3	24	72

b) Byer mindst 100.000 indbyggere

		2021										
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	
1992	D1	[0]	1	6	16	27	20	22	8	1	0	0
	D2	[2]	0	1	7	21	29	23	17	2	0	0
	D3	[5]	0	0	1	11	32	29	21	5	1	0
	D4	[10]	0	0	0	4	20	34	30	10	2	0
	D5	[14]	0	0	0	1	10	31	37	18	4	0
	D6	[15]	0	0	0	0	3	20	38	27	10	1
	D7	[14]	0	0	0	0	1	9	29	34	25	3
	D8	[14]	0	0	0	0	0	5	19	29	38	9
	D9	[16]	0	0	0	0	0	2	13	27	37	21
	D10	[17]	0	0	0	0	0	0	3	16	32	49

c) Byer 50.000-99.999 indbyggere

		2021										
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	
1992	D1	[2]	5	36	27	20	9	3	0	0	0	0
	D2	[7]	1	8	24	32	23	9	2	0	0	0
	D3	[11]	0	1	11	32	38	15	4	0	0	0
	D4	[13]	0	0	3	26	43	21	6	1	0	0
	D5	[13]	0	0	1	14	44	32	8	1	0	0
	D6	[11]	0	0	0	5	30	45	16	4	1	0
	D7	[9]	0	0	0	2	14	43	23	13	4	0
	D8	[10]	0	0	0	1	9	34	27	15	12	1
	D9	[10]	0	0	0	0	4	24	29	15	25	3
	D10	[6]	0	0	0	0	0	7	12	11	52	19

d) Byer 20.000-49.999 indbyggere

		2021										
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	
1992	D1	[1]	12	22	28	21	12	4	1	0	0	0
	D2	[6]	4	14	27	31	16	7	2	0	0	0
	D3	[12]	1	12	23	28	21	11	4	1	0	0
	D4	[18]	0	6	16	27	24	15	11	2	0	0
	D5	[22]	0	1	9	21	27	19	17	5	1	0
	D6	[22]	0	0	3	10	22	21	26	16	2	0
	D7	[20]	0	0	2	5	11	17	23	35	8	0
	D8	[22]	0	0	1	3	7	14	16	42	17	0
	D9	[26]	0	0	0	1	4	14	16	32	32	2
	D10	[23]	0	0	0	0	0	3	8	21	48	20

e) Byer 1.000-19.999 indbyggere

		2021										
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	
1992	D1	[34]	53	31	12	3	1	0	0	0	0	0
	D2	[60]	22	35	25	11	4	1	0	0	0	0
	D3	[68]	7	25	30	20	10	5	1	0	0	0
	D4	[64]	2	14	25	25	16	13	5	1	0	0
	D5	[56]	0	6	16	23	19	19	14	2	0	0
	D6	[51]	0	3	8	16	20	21	24	7	1	0
	D7	[47]	0	1	4	11	17	21	26	16	2	0
	D8	[40]	0	0	2	7	13	19	26	25	8	0
	D9	[30]	0	0	0	2	6	12	24	36	18	2
	D10	[19]	0	0	0	0	1	2	8	32	40	16

f) Byer færre end 1.000 indbyggere

		2021										
		D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	
1992	D1	[101]	70	22	6	2	0	0	0	0	0	0
	D2	[63]	32	36	20	8	3	1	0	0	0	0
	D3	[40]	13	30	29	17	8	3	1	0	0	0
	D4	[29]	6	20	27	24	14	7	2	0	0	0
	D5	[24]	2	13	21	24	19	12	7	1	0	0
	D6	[21]	1	9	17	23	22	15	11	3	0	0
	D7	[18]	0	5	12	20	22	18	15	6	1	0
	D8	[13]	0	3	8	15	20	22	18	11	3	0
	D9	[10]	0	2	6	10	16	20	23	17	7	1
	D10	[8]	0	1	3	6	9	13	20	24	20	4

Anm.: På landsplan er alle ejerboliger rangeret efter boligens skønnede kvadratmeterpris og inddelt i deciler. Første decil D1 indeholder boligerne med lavest kvadratmeterpris. Figuren viser prisdecil for boligen i 1992 (nedad) og i 2021 (udad). Bystørrelse opgøres ultimo 2020.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

I hovedstadsområdet og for byer med mere end 100.000 indbyggere er ejerboliger allerede i 1992 relativt dyre og typisk at finde i femte decil eller højere. Alligevel ses en markant opadgående prisbevægelse fra 1992 til 2021 (en bevægelse mod højre i prismobilitetsmatricen).

Uanset hvilken decil ejerboligerne i hovedstadsområdet tilhørte i 1992 er over 70 pct. af ejerboligerne i 2021 placeret blandt de 20 pct. dyreste boliger i hele landet målt på kvadratmeterpris.

For byer med mere end 100.000 indbyggere ses en tendens til at ejerboligerne er flyttet 2-3 deciler op fra 1992 til 2021. Mens byer med 50.000-99.999 er mere diagonalt centreret men dog generelt flyttet et decil op. I byer med mindre end 50.000 indbyggere ses en diagonalt centreret tendens, hvor prismobilitet går op og ned med et decil i forhold til udgangspunktet i 1992. De mindste byområder med færre end 1.000 personer og landdistriktet har en nedadgående tendens hvor især ejerboliger, der tilhørte den dyreste halvdel af boligerne i 1992 er faldet 2-3 deciler i 2021.

#### Boks 4.2

##### Kullback-Leibler divergens

Kullback-Leibler divergens sammenholder en faktisk spatial fordeling med en teoretisk fordeling, hvor punkterne er fuldstændig tilfældigt placeret (eng: complete spatial randomness, CSR).

Landet inddeles i  $N$  unikke områder, som er konstante gennem den betragtede tidsperiode. Konkret inddeles i hektarceller.

Lad  $x_{i,t,d}$  være antallet af ejerboliger i område  $i$  ved tidspunkt  $t$ , der tilhører decilgruppe  $d$ . Ud fra data beregnes den faktiske sandsynlighed for, at en ejerbolig fra prisdecil  $d$  på tidspunkt  $t$  er beliggende i område  $i$ :

$$p_{i,t,d} = \frac{x_{i,t,d}}{X_{t,d}} \quad \text{hvor } X_{t,d} = \sum_{i=1}^N x_{i,t,d}$$

Den teoretiske fordeling er karakteriseret af homogenitet og uafhængighed, dvs. at intet punkt har præference for en given spatial placering, og at informationen omkring udfaldet i et område ikke påvirker udfaldet i et andet område. Lad  $a_i$  være arealet af område  $i$ . Den teoretiske sandsynlighed, som er konstant over tid, er givet ved:

$$p_i^{CSR} = \frac{a_i}{A} \quad \text{hvor } A = \sum_{i=1}^N a_i$$

For diskrete sandsynlighedsfordelinger er Kullback-Leibler divergensen fra den faktiske observerede til den teoretiske sandsynlighedsfordeling defineret som:

$$KL_{t,d} = \sum_{i=1}^N p_{i,t,d} \cdot \log\left(\frac{p_{i,t,d}}{p_i^{CSR}}\right)$$

Hvis  $KL_{t,d} = 0$  er de to sandsynlighedsfordelinger ens. I så fald er ejerboligerne i prisdecil  $d$  tilfældigt spatialt fordelt ud over hele Danmark på tidspunkt  $t$ .

Hvis divergensindekset for to prisdeciler er lig hinanden, vil ejerboligerne for disse to grupper være sammenklynget i enslydende grad. Hvis  $KL_{t,D1} > KL_{t,D2}$ ,  $D1 \neq D2$  betyder det, at prisdecil  $D1$  ved tidspunkt  $t$  er mere sammenklynget end prisdecil  $D2$ .

## Geografisk koncentration

Af de foregående delafsnit fremgår det, at ikke kun placeringen af prisområderne har flytte sig fra 1992 til 2021, men også, at den geografiske koncentration, rumlig adskillelse eller sammenklynkning af ejerboliger inden for forskellige prisdeciler har ændret sig i perioden.

Til at måle udviklingen i den geografiske koncentration af ejerboliger inden for en kvadratmeterpris-decil, opstilles det såkaldte Kullback-Leibler divergens mål, jf. Kullback og Leibler (1951). Målet er også kendt som relativ entropi. Metoden er nærmere beskrevet i Boks 4.2 på foregående side.

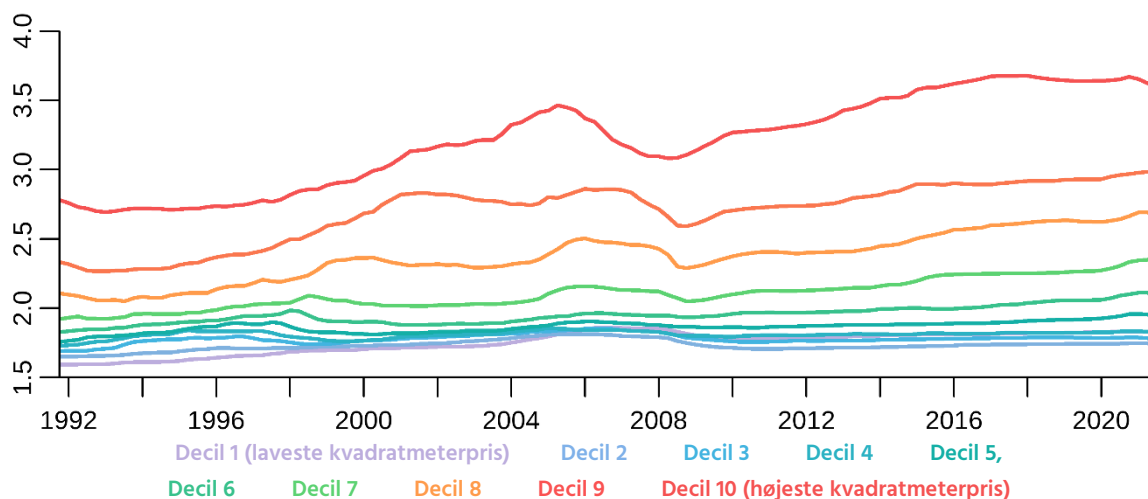
Hvis indekset er nul, er ejerboligerne i den pågældende decil tilfældigt fordelt geografisk, og der er ingen tegn på klyngedannelse. Jo større indeks værdi, desto større er den rumlige klyngedannelse af ejerboliger i det pågældende decil og jo større er graden af boligsegregering.

Generelt gælder det, at jo højere kvadratmeterprisen er på ejerboliger, desto mere sammenklynget er boligerne geografisk, jf. Figur 4.5. For den dyreste halvdel af boligerne målt på kvadratmeterpris ses en klar hierarkisk opdeling efter deciler. Der ses ikke samme grad af klyngedannelse blandt den billigste halvdel af ejerboliger.

Gennem perioden har den geografiske koncentration af relativt billige ejerboliger pr. kvadratmeter være nogenlunde konstant. De dyreste ejerboliger bliver tættere geografisk koncentreret over de seneste 30 år.

I perioden op til finanskrisen i 2007 ses det, at de 30 pct. dyreste ejerboliger bliver relativt mere sammenklynget end resten af ejerboligmarkedet. Den samme tendens ses i opgangsperioden fra 2013 og frem.

**Figur 4.5**  
Geografisk koncentration målt ved Kullback-Leibler divergens



Anm.: Alle ejerboliger er rangeret i deciler efter boligens skønnede kvadratmeterpris. Figuren viser, hvor geografisk koncentreret boligerne i hvert prisdecil er. Hvis værdien er nul, er ejerboligerne i det pågældende decil tilfældigt fordelt geografisk

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

## 4.2 Geografisk variation i årlig vækstrate i boligpriserne

I det følgende undersøges det hvilke geografiske områder, der i perioden siden 1992 har oplevet store prisstigninger og hvilke dele af landet, som ikke i samme grad har oplevet en opadgående pris. Herved illustreres det, i hvilke dele af landet boligprisen trak fra andre egne.

Udviklingen illustreres på landkort. Derfor betragtes boligprisen for en standard ejerbolig med den unikke geografiske beliggenhed som faktisk observeret.

Undersøgelsen sker ved over en periode at udregne den gennemsnitlige årlige procentvise ændring i prisen pr. boligkvadratmeter. Prisændringen inddeles i intervaller. Periodelængden vil være hele dataperioden eller relevante delperioder. Beregningen sker for hver ejerbolig.

På et landkort sættes en prik ved boligens geografiske placering. Prikken farvelægges efter, hvilken intervalkategori prisændringen for netop denne ejerbolig tilhører. Farvekoden for prikkerne på kortet afspejler således det gennemsnitlige årlige kapitalafkast som boligejeren har opnået gennem det betragtede periode.

I det følgende betragtes først Danmark som helhed, hvorefter de fem regioner beskrives selvstændigt. Dog opdeles region Syddanmark i de to landsdele Fyn og Syddanmark. Region Sjælland og region Hovedstaden vises som ét samlet område, idet boligmarkedet i disse områder er tæt forbundne. For Danmark som helhed beskrives desuden udviklingen i to delperioder omfattende opturen på boligmarkedet i henholdsvis 1999-2006 og 2013-2021. For de enkelte regioner er tilsvarende figurer medtaget i appendiks B.

### Udviklingen i Danmark som helhed

Figur 4.6 viser den skønnede gennemsnitlige årlige ændring i boligprisen pr. kvadratmeter over den 30-årige periode 1992-2021. Boligerne er inddelt i grupper efter afkastets størrelse. Farvekoden for prikkerne på kortet indikerer, hvilken gruppe den enkelte bolig tilhører. Det gennemsnitlige årlige afkast på ejerboligen vurderes at variere mellem knap 2 pct. årligt og op til over 8 pct. årligt.

Til sammenligning har den almindelige prisstigningstakt i perioden være 1,7 pct. årligt<sup>8</sup> og aktier steg gennemsnitligt med omkring 8-9 pct. årligt fra primo 1992 til ultimo 2021<sup>9</sup>.

Overordnet kan man opdele landet i tre områder: Et højt afkast på mindst 6 pct. årligt. Dette omfatter de røde og orange områder på kortet. Disse boliger er særligt koncentreret omkring København, men der ses også områder i eksempelvis Aarhus, Aalborg, Odense og Roskilde i denne gruppe. I det centrale København er den gennemsnitlige prisstigning over 8 pct. årligt. For Aarhus er stigningen oftest i niveauet 6½-7 pct., mens Aalborg, Odense og Roskilde ligger på 5½-6 pct. og kun visse dele af disse byområder når over 6 pct.

For det andet et middel afkast på 4-6 pct. årligt, hvilke omfatter områder i grønne nuancer på kortet. Dette omfatter dels centralt beliggende boliger i de større provinsbyer som eksempelvis Silkeborg, Skanderborg, Vejle, Horsens, Esbjerg, Svendborg, Holbæk med flere og dels oplandet til de største byer. Her aftager prisstigningen aftages typisk, når afstanden til bykernen øges. Dog med en vis variation på tværs af landet, hvilket beskrives nærmere ved beskrivelsen af de enkelte delområder nedenfor. Særligt omkring København ses et stort område i denne kategori. Området dækker store dele af Nordsjælland, ud omkring Roskilde i vest og ned til Køge og det nordligste af Stevn i syd.

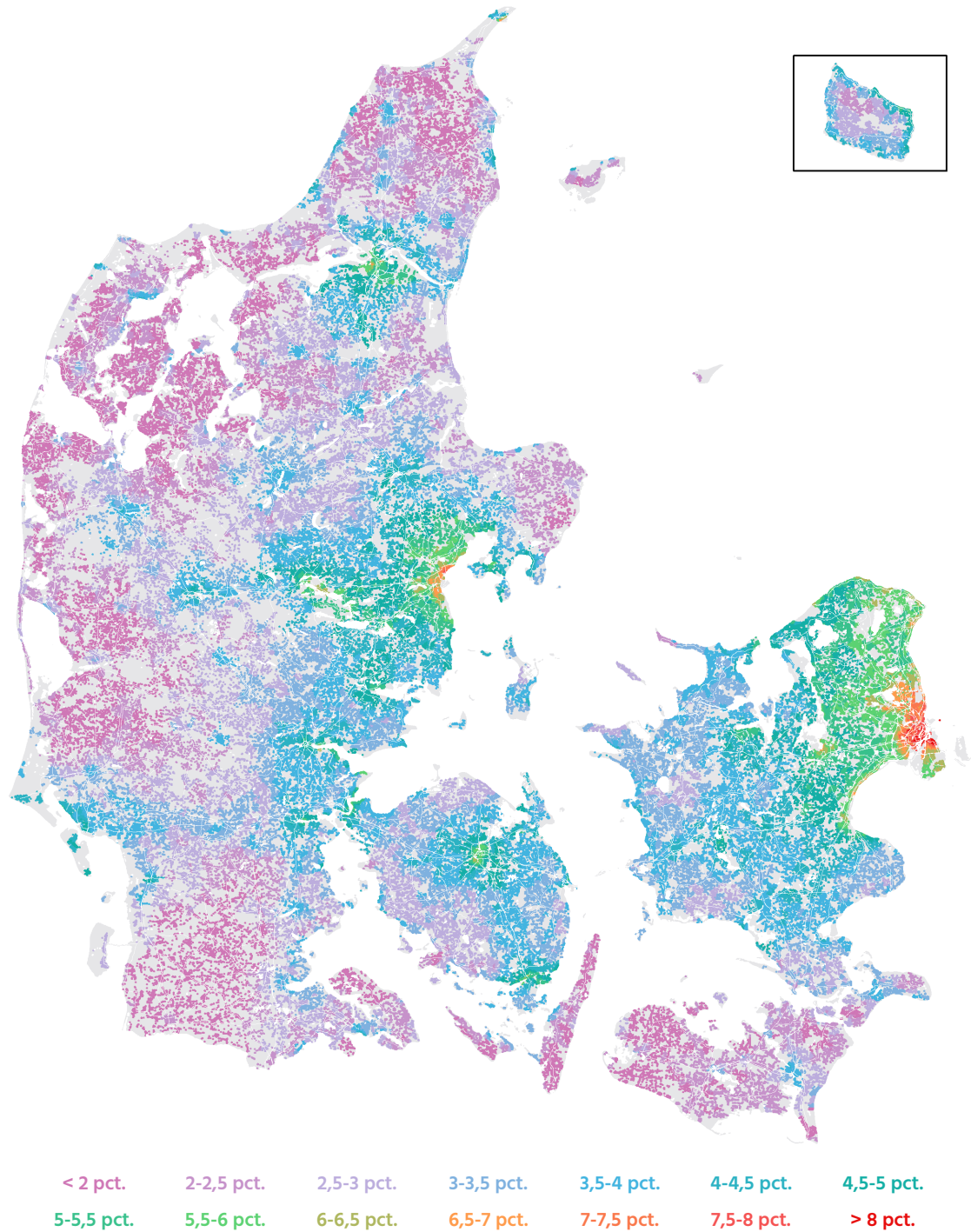
<sup>8</sup> Ifølge Danmarks Statistiks forbrugerprisindeks.

<sup>9</sup> Det amerikanske aktieindeks S&P500 steg med i gennemsnit 8,5 pct. årligt i perioden 1992- 2021. Det danske aktieindeks OMXC20 steg med 9 pct. årligt i perioden medio 1989 til ultimo 2021, jf. [www.statistikbanken.dk/MPK13](http://www.statistikbanken.dk/MPK13). OMXC20 er et indeks bestående af de 20 største og mest handlede aktier noteret på Københavns Fondsbørs.

Sidst et stort geografisk område med et relativt lavt afkast på under 4 pct. årligt. Dette er de blå og lilla områder på kortet. Dette omfatter især de mindre byer og landdistrikterne. Eksempler på disse områder er Nordjylland foruden Aalborg, Vestjylland, store dele af Fyn, Lolland-Falster og det sydvestlige Sjælland.

**Figur 4.6**

Gennemsnitlig årlig ændring i boligprisen pr. kvadratmeter, 1992-2021



Anm.: Figuren viser den estimerede gennemsnitlige årlige procentvise ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

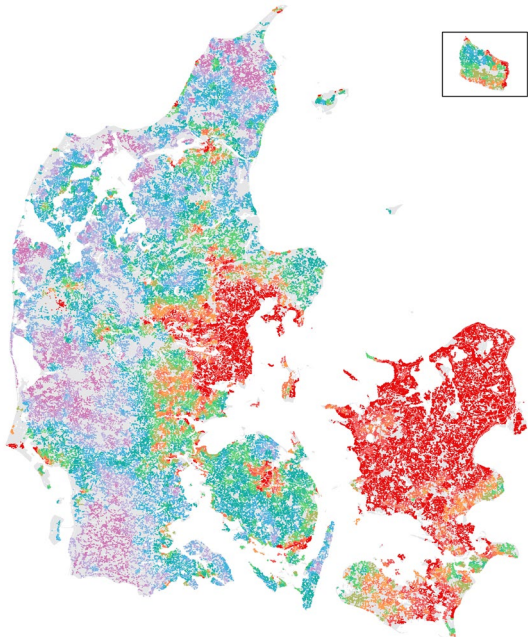
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.



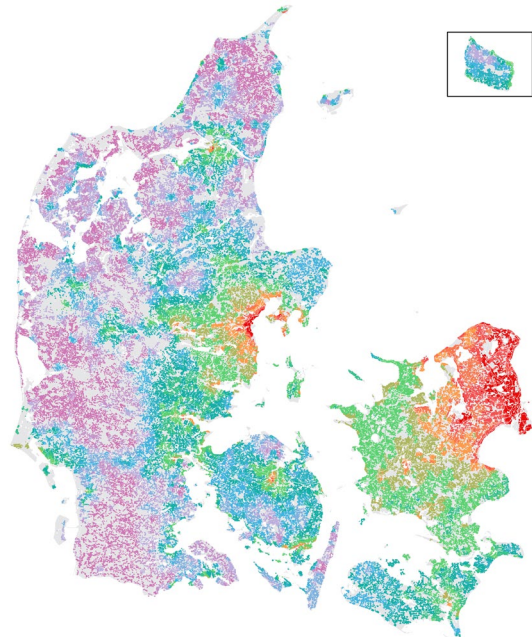
Figur 4.7

Gennemsnitlig årlig ændring i boligprisen pr. kvadratmeter, 1999-2006 og 2013-2021

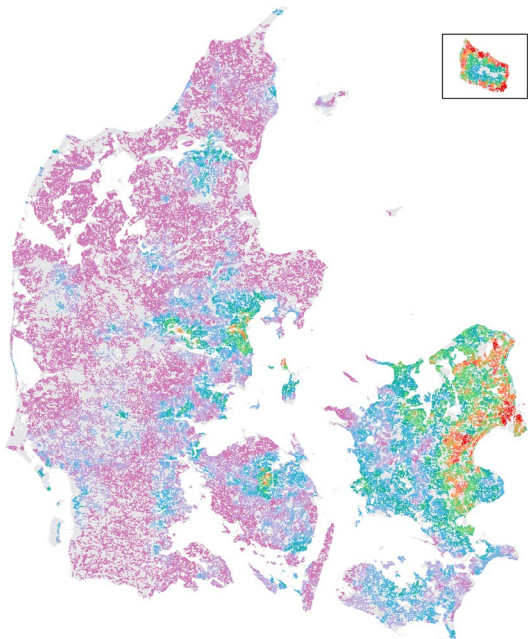
a) Gns. årlig ændring 1999-2006, pct.



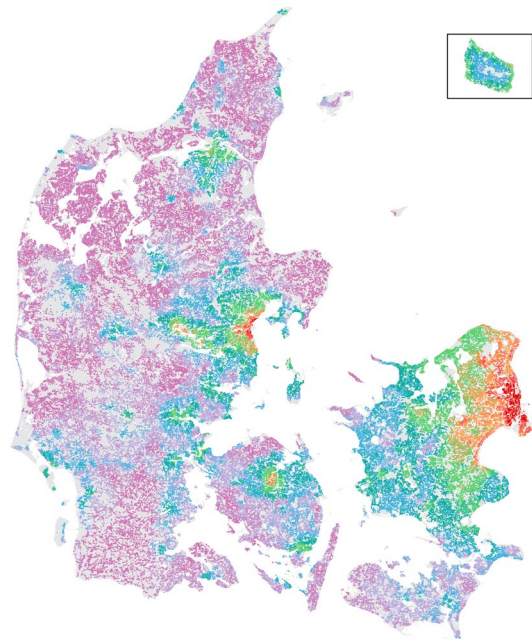
b) Gns. årlig ændring 1999-2006, kr. (2021-niveau)



c) Gns. årlig ændring 2013-2021, pct.



d) Gns. årlig ændring 2013-2021, kr. (2021-niveau)



< 2 pct.	2-2,5 pct.	2,5-3 pct.	3-3,5 pct.	< 50 kr.	50-100	100-150	150-200
3,5-4 pct.	4-4,5 pct.	4,5-5 pct.	5-5,5 pct.	200-250	250-300	300-400	400-500
5,5-6 pct.	6-6,5 pct.	6,5-7 pct.	7-7,5 pct.	500-750	750-1.000	1.000-1.250	1.250-1.500
7,5-8 pct.	> 8 pct.			1.500-2.000	> 2.000 kr.		

Anm.: Figuren viser den estimerede gennemsnitlige årlige procentvise ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

## Udviklingen i Danmark som helhed: Sammenligning af opgangsperioder

Følgende afsnit fokuserer på den geografiske udvikling i ejerboligpriserne for de to delperioder, hvor priserne steg mest markant. Dette er årene op til finanskrisen, som afgrænses til årene 1999-2006, og perioden hvor økonomien igen var kommet sig over krisen, afgrænset til årene 2013-2021.

I begge delperioder steg boligprisen pr. kvadratmeter over hele landet, jf. Figur 4.7 på foregående side. I den første delperiode 1999-2006 steg priserne markant mere, og de kraftige prisstigninger dækkede et større geografisk område, end de gjorde i perioden 2013-2021. På stort set hele Sjælland og omkring de største byer, herunder især Aarhus, steg priserne gennemsnitligt med over 8 pct. årligt, hvilket i perioden også var tilfældet for dele af Lolland, Falster og Sydfyn. I Aarhus, København og Nordsjælland svarede væksten til en gennemsnitlig stigning i kvadratmeterprisen på over 2.000 kr. pr. år. På Lolland, Falster og Sydfyn svarede til en gennemsnitlig årlig stigning i kvadratmeterpriserne på omkring 250-500 kr.

I anden delperiode var prisstigningerne mere lokalt forankret i og tæt omkring de større byer. Her var det få områder, som for eksempel omkring Hovedstaden og på Bornholm, der oplevede samme vækst i ejerboligpriserne som op til finanskrisen. I områderne omkring Aarhus, Nordsjælland og Sydfyn var prisstigningerne i denne periode på under 4 pct. årligt, hvilket også gjorde sig gældende for Lolland og Falster.

Bemærk at ejerboligerne i Nordsjælland kom fra en relativ høj kvadratmeterpris i 2013 så prisstigningen på omkring 4 pct. årligt svarede til en gennemsnitlig årlig stigning i kvadratmeterpriserne på omkring 2.000 kr. (2021-niveau). Hvorimod ejerboligerne på eksempelvis Lolland og Falster kom fra et relativt lavt prisniveau, hvorved væksten her svarede til en gennemsnitlig stigning i kvadratmeterpriserne på omkring 100 kr. pr. år (2021-niveau).

## Region Hovedstaden og region Sjælland

Siden 1992 er boligprisen pr. kvadratmeter øget forholdsmeæssigt mest i det centrale København, hvor prisen i gennemsnit er steget med mere end 8 pct. årligt, jf. Figur 4.6. Området omfatter indre by, brokvartererne, Christianshavn samt dele af Amager, særligt området ved Islands Brygge samt Amager Strand.

Fra det centrale København aftager prisstigningerne, når afstanden til indre by øges. Der ses tendens til højere prisstigning i området umiddelbart nord for København (orange svarende til gennemsnitlig årlig prisstigning på 6½-7 pct.) i forhold til områder med samme afstand til indre by beliggende vest og sydvest for byen (grøn 5-5½ pct.).

Prisstigningerne i kystnære områder i Hovedstadsområdet er generelt højere end i områder med større afstand til kysten. Dette gælder eksempelvis langs Køge bugt og langs den nordsjællandske kyststrækning.

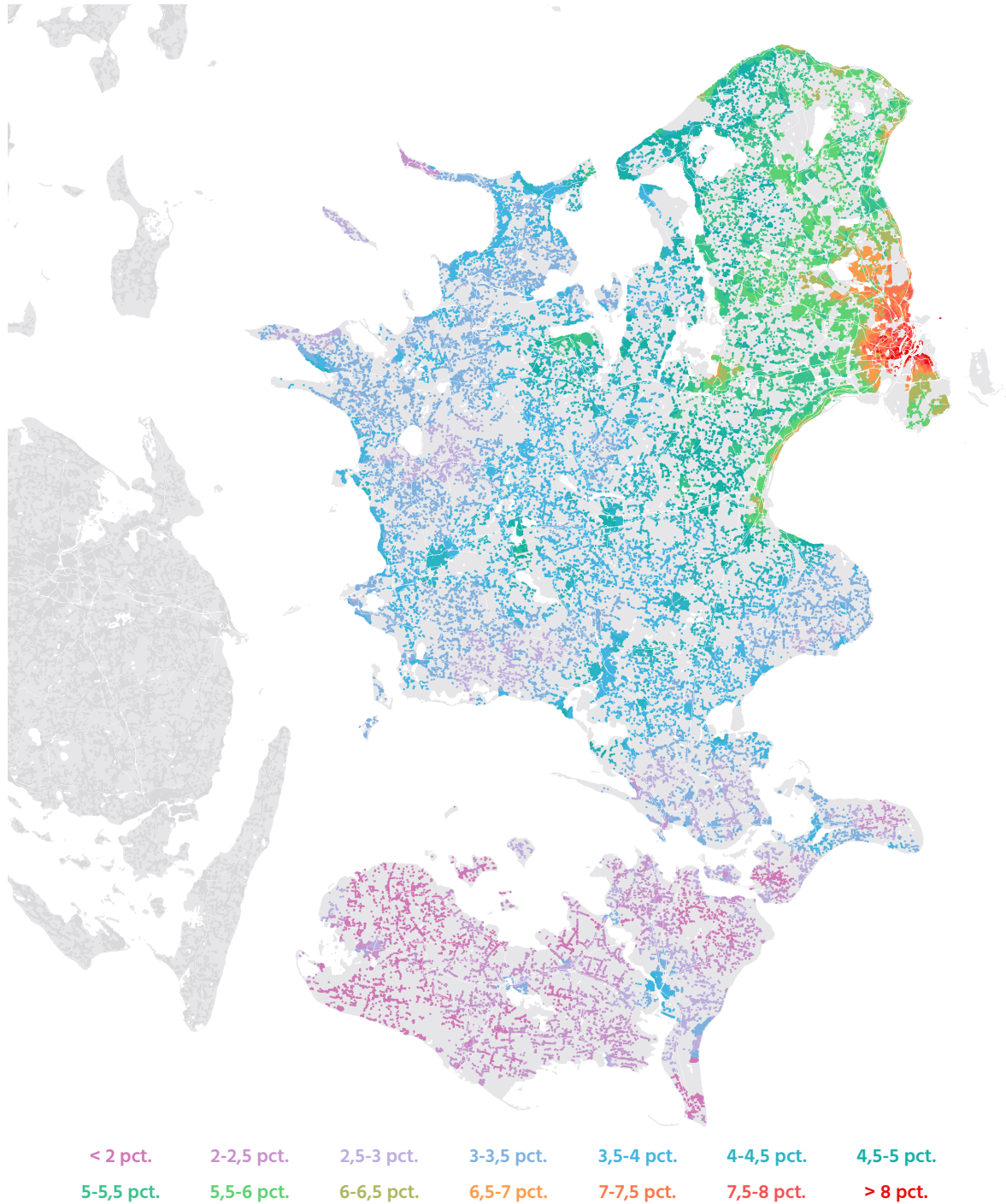
Der ses også en relativt høj årlig vækst i boligprisen pr. kvadratmeter i de større byer uden for København. Dette gælder de grønne områder i eksempelvis Holbæk, Ringsted, Sorø, Slagelse og Næstved. Inden for disse byer ses ligeledes variation i prisudviklingen.

Det er værd at bemærke, at der ses tendens til højere vækst i boligprisen langs den øst-vestgående motorvej E20, som forbinder Hovedstadsområdet med Fyn og Jylland. Denne tendens ses ikke i samme grad langs den sydgående motorvej E47, som forbinder hovedstadsområdet med Tyskland via Rødby.

Relativt lav årlig vækst i ejerboligprisen ses i flere større områder. Områderne inkluderer Lolland og Falster, det sydlige Stevns og det vestlige samt sydlige Sjælland. Inden for disse områder ses dog også delområder med højere vækst, hvilket eksempelvis gælder Nykøbing Falster, Marielyst, dele af Møn og Kalundborg. En del af områderne med relativt lav vækst er halvøer, andre af disse områder er karakteriseret ved, at de ligger langt fra de større byer.

**Figur 4.8**

Gennemsnitlig årlig ændring i boligprisen pr. kvadratmeter, region Hovedstaden og Sjælland, 1992-2021



Anm.: Figuren viser den estimerede gennemsnitlige årlige procentvise ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

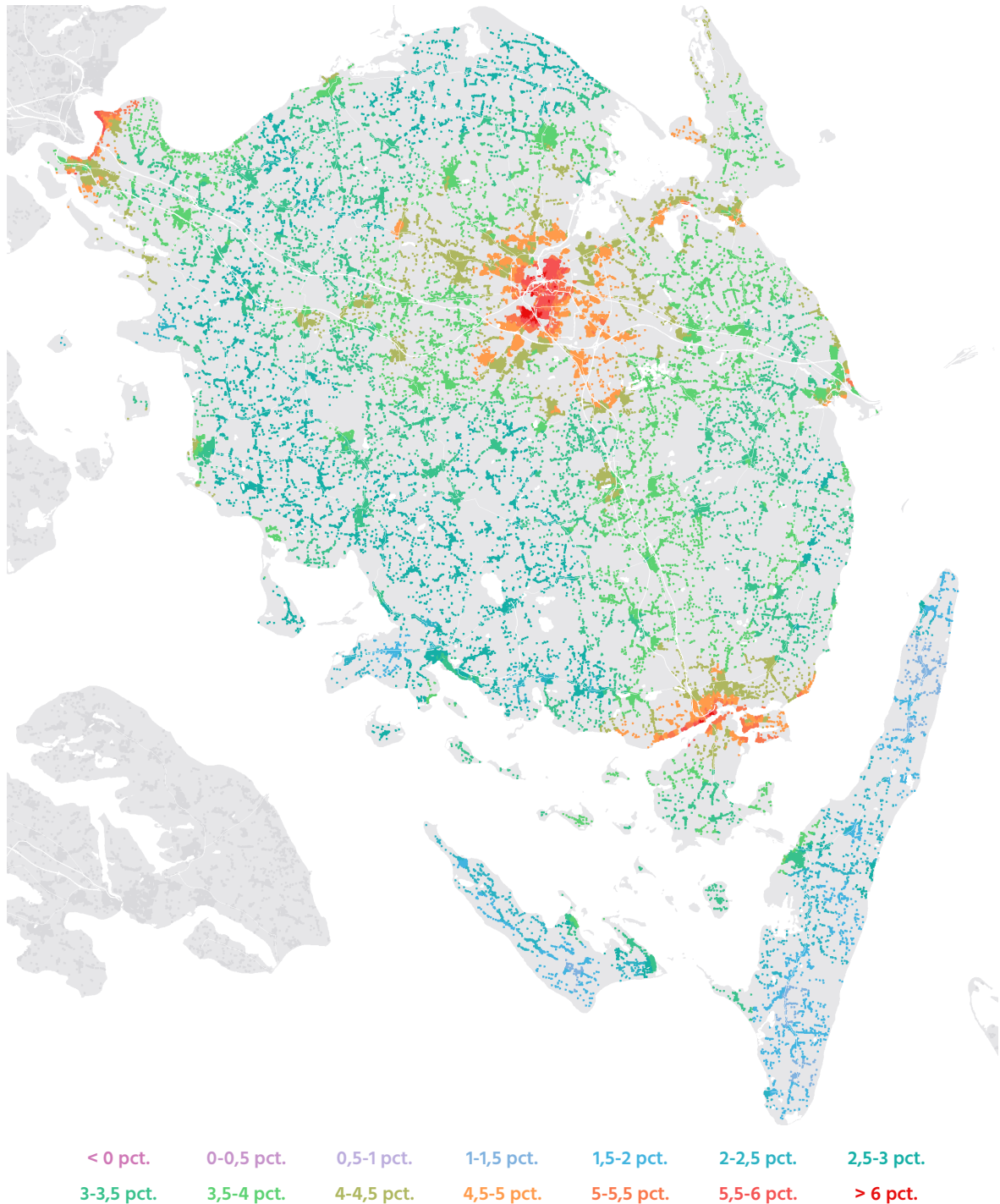
### Landsdel Fyn

På Fyn vurderes de største prisstigninger i det centrale Odense, hvor kvadratmeterprisen estimeres at stige med over 6 pct. årligt i perioden 1992-2021, jf. Figur 4.9 (de røde områder). Fra det centrale Odense reduceres prisstigningstakten, når afstanden til bycentrum øges.

Også Svendborg inkl. Vindeby og Troense på Tåsinge, Middelfart og Munkebo oplever pæne årlige prisstigninger i niveauet 4½-5 pct. årligt (de orange områder). I disse byer ses områder langs kysten, som vurderes at stige mere end byen generelt.

Figur 4.9

Gennemsnitlig årlig ændring i boligprisen pr. kvadratmeter, landsdel Fyn, 1992-2021



Anm.: Figuren viser den estimerede gennemsnitlige årlige procentvise ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

I de øvrige større byer som eksempelvis Nyborg, Assens, Bogense, Kerteminde og til dels også Faaborg og Rudkøbing ses årlige prisstigninger i niveauet cirka 2 pct. årligt (grønne områder). Dette er højere end de tyndere befolkede områder umiddelbart omkring byerne. Samme tendens ses på øerne, hvor byområder vurderes at have større årlige prisstigninger end landdistrikterne. Dette gælder for eksempel Rudkøbing på Langeland eller Ærøskøbing og Marstal på Ærø.

På Fyn ses desuden tendens til en relativt kraftig stigning i boligprisen omkring de centrale transportkorridorer sammenlignet med øvrige områder i samme afstand til eksempelvis Odense. Dette gælder for eksempel langs den øst-vest gående motorvej på tværs af Fyn og langs den sydgående motorvej fra Odense til Svendborg.

### Landsdel Sydjylland

I Sønderjylland vurderes den årlige stigning i kvadratmeterprisen for en ejerbolig at være kraftigst i Trekantsområdet, jf. Figur 4.10. Prisstigningen er særligt stor i byområderne i form af Kolding, Vejle og Fredericia. Her vurderes prisstigningen til over 5 pct. årligt (de orange-røde områder). For alle tre byer gælder, at prisudviklingen vurderes kraftigere i de kystnære områder end i områder længere fra vandet. Også i Esbjerg vurderes prisudviklingen at være relativt høj, her vurderes priserne at stige med 4-5 pct. årligt (grønne-orange områder). Også langs den øst-vest gående motorvej, som forbinder Trekantsområdet og Esbjerg skønnes en relativt kraftig prisudvikling sammenlignet med tilsvarende områder længere fra motorvejen.

I de øvrige større byer som eksempelvis Haderslev, Aabenraa, Sønderborg, Ribe, Varde, Billund og Give vurderes en pæn vækst i kvadratmeterpriserne i niveauet 2-3 pct. årligt. I yderområderne og på øerne estimeres prisudviklingen til at være mere afdæmpet.

### Region Midtjylland

I region Midtjylland vurderes den årlige stigningstakt i det centrale Aarhus til over 7 pct. årligt i perioden 1992-2021, jf. Figur 4.11 (de røde områder). Fra bymidten aftages prisstigningstakten, når afstanden til centrum øges. Der ses tendens til større prisstigninger nord for byen samt langs kysten syd for Aarhus end i områder af tilsvarende afstand fra bymidten nordvest samt sydvest for byen.

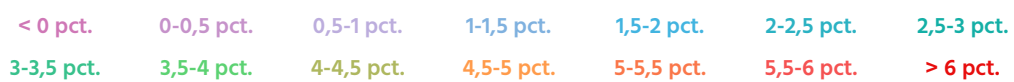
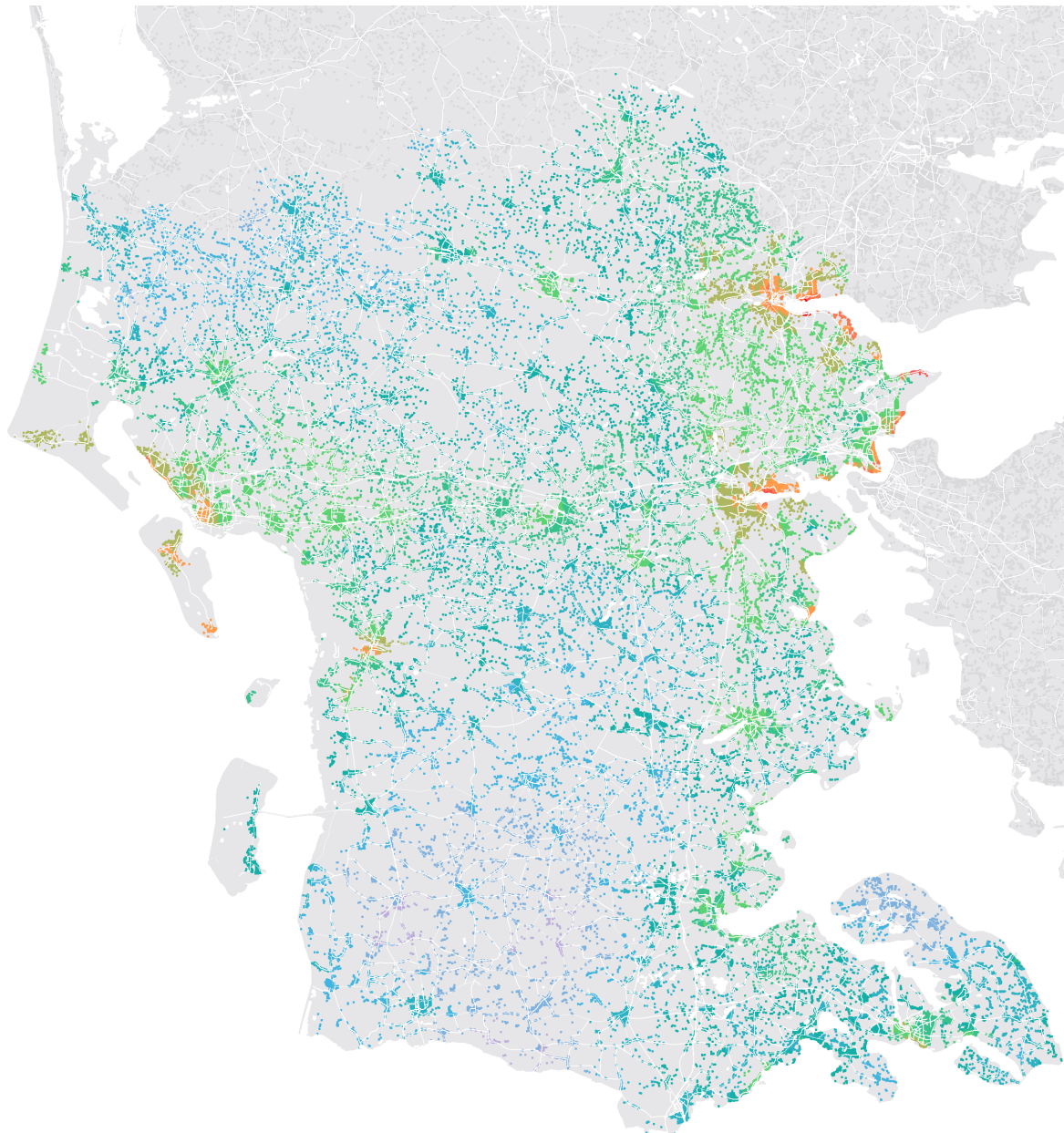
Der ses også en relativt kraftig vækst i boligpriserne i et bælte fra Aarhus og mod vest forbi Skanderborg og Silkeborg og til dels også videre mod Ikast og Herning. Også mod syd og ned omkring Horsens vurderes en relativt høj prisstigning, dog lavere end i Skanderborg og Silkeborg. Også nord og øst for Aarhus ses en relativt høj prisstigning i og omkring henholdsvis Randers og Ebeltoft, hvor prisstigningen er på niveau med Horsens.

Randers, Horsens og Herning er større byer end Silkeborg og Skanderborg, men førstnævnte oplever alligevel lavere vækst i boligprisen end sidstnævnte. Årsagen er sandsynligvis, at prisniveauet i Skanderborg og Silkeborg trækkes op af nærheden til Aarhus samt en naturskøn beliggenhed.

I den øvrige del af regionen skiller de større byer som Viborg, Holstebro og Brande sig ud med relativt høj prisstigningstakt sammenlignet med sine umiddelbare nærområder. Det samme gør sig gældende for eksempelvis Ringkøbing, Videbæk, Lemvig, Skive og Grenaa, hvor prisstigningstakten er større end de omkringliggende landdistrikter. I landdistrikterne vurderes kvadratmeterprisen at være steget med mindre end 1 pct. årligt (lilla områder).

Figur 4.10

Gennemsnitlig årlig ændring i boligprisen pr. kvadratmeter, landsdel Sydjylland, 1992-2021

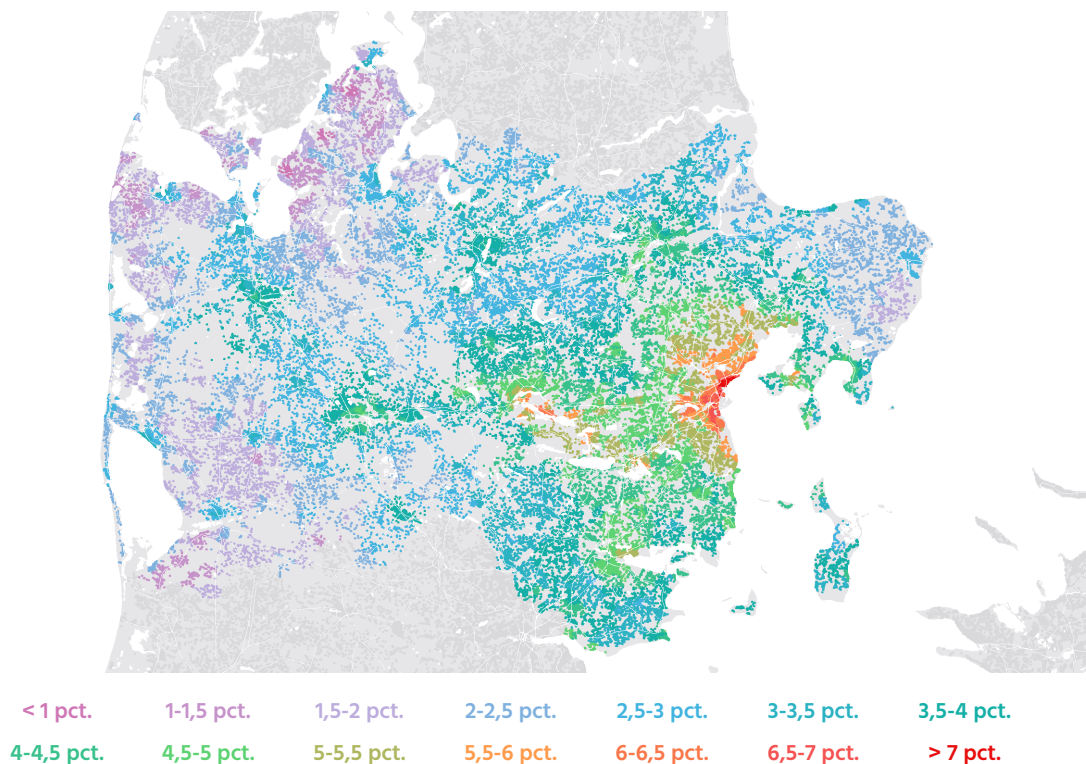


Anm.: Figuren viser den estimerede gennemsnitlige årlige procentvise ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur 4.11

Gennemsnitlig årlig ændring i boligprisen pr. kvadratmeter, region Midtjylland, 1992-2021



Anm.: Figuren viser den estimerede gennemsnitlige årlige procentvise ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

### Region Nordjylland

I region Nordjylland sker den største årlige stigning i kvadratmeterprisen i det centrale Aalborg, særligt i byområdet vest for centrum, jf. Figur 4.12. Her skønnes kvadratmeterprisen at stige med mere end 6 pct. årligt i perioden 1992-2021 (røde områder), hvilket er i samme størrelsesorden som i Odense.

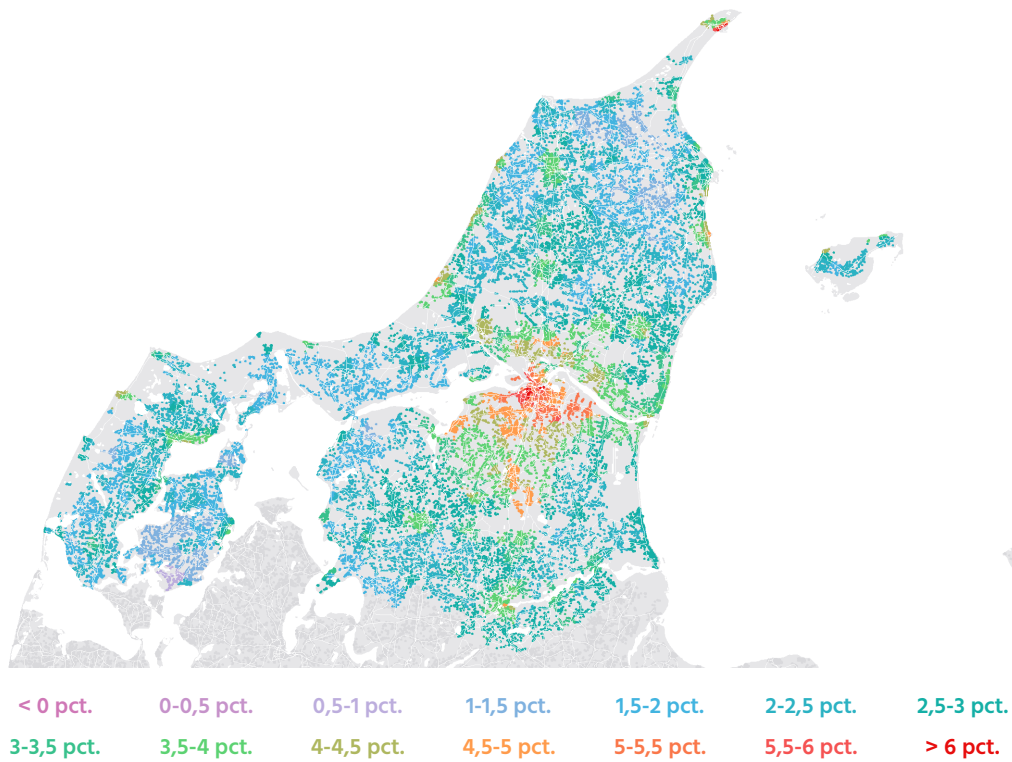
I Nordjylland ses tendens til opdeling i fire områder. Det første er som nævnt Aalborg, Nørresundby og opland, hvor den årlige prisstigning skønnes til over 3 pct. årligt (røde, orange og grønne områder). Syd for Limfjorden vurderes at større område i oplandet til Aalborg at have oplevet en relativt stor prisstigning, end det er tilfældet i samme afstand nord for fjorden.

Det andet område er de større byer eksempelvis Thisted, Hjørring, Hobro, Sæby og Skagen. Her vurderes en relativ kraftig prisstigning.

I det resterende område skønnes prisstigningen mere begrænset. Området kan dog alligevel deles i to, idet visse mindre byområder såsom Klitmøller, Aars og langs kysten ved Blokhus, Løkken samt Lønstrup ligeledes oplever pæne prisstigninger. For sidstnævnte kystområde kan det spille ind, at visse huse her kan anvendes som feriehuse, hvilket kan have trukket prisen i opadgående retning. De mindre byer og landdistrikterne har typisk oplevet en relativt svagt prisstigning.

**Figur 4.12**

Gennemsnitlig årlig ændring i boligprisen pr. kvadratmeter, region Nordjylland, 1992-2021



Anm.: Figuren viser den estimerede gennemsnitlige årlige procentvise ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.







## 5. Metode til vurdering af boligprisudviklingen

For hver enkelt ejerbolig i Danmark skønnes ejendomsværdien i perioden 1992-2021. Værdien fastsættes ud fra geografisk vægtning af handlede boliger i nærområdet til den betragtede bolig. Sammenhængen mellem boligens karakteristika og kvadratmeterprisen fastsættes ved en nutidig metode, der baseres på maskinlæring (eng: machine learning).

For samtlige 1,4 mio. privatejede ejerboliger vurderes ejendomsværdien i hvert kvartal i perioden 1992-2021. Dette gælder uanset, om boligen i løbet af perioden er handlet eller ej. Den anvendte metode har tre væsentlige karakteristika.

For det første foretages en selvstændig prisvurdering for hver enkelt ejerbolig. Herved kan prisseffekten af de indeholdte boligkarakteristika variere på tværs af landet, så for eksempel betydningen af en beliggenhed tæt ved vandet er anderledes på en ø end i byen. I byen vil der typisk være flere potentielle købere til relativt få boliger ved vandet, hvorfor merprisen ved en kystnær bolig kan være højere end i øvrige dele af landet.

For det andet anvender metoden geografisk vægtning. Prisen på en bolig fastsættes ud fra handlede boliger i nærområdet til boligen selv. Hver handlet bolig vægtes ud fra afstanden til den betragtede bolig. Bolighandler tæt ved vægtes højt, og ejendomssalg længere væk vægtes lavere. Den geografiske vægtning giver en fin balance mellem lokale og globale egenskaber i beskrivelsen af boligprisen.

For det tredje anvendes en algoritme baseret på maskinlæring til at fastsætte sammenhængen mellem boligens karakteristika og kvadratmeterprisen. Algoritmen er udvalgt efter test af flere forskellige metoder. Den her anvendte algoritme giver således en mere præcis vurdering af boligpriserne end simpel regression, som er anvendt i en lignende analyse foretaget af DREAM i 2018.

Afsnit 5.1 indeholder et illustrativt eksempel på den anvendte metode. I den resterende del af kapitlet gives en gennemgang af metoden målrettet læsere med en basal viden inden for statistiske og økonomiske metoder.

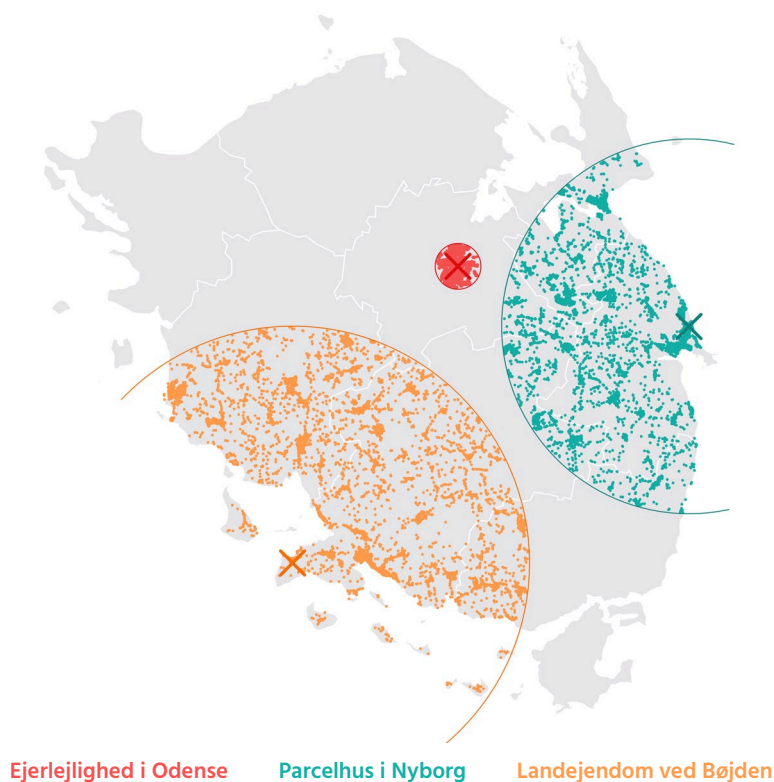
### 5.1 Illustrativt eksempel på geografisk vægtet regression

I det følgende illustreres principperne bag geografisk vægtet regression ved et konkret eksempel. Der tages udgangspunkt i tre ejerboliger beliggende på Fyn.

Den ene bolig er en mindre ejerlejlighed i Odense på 38 kvadratmeter i Odense. Den anden bolig har adresse i Nyborg og er et parcelhus på 160 kvadratmeter opført i 1970. Den tredje og sidste bolig er en landejendom på 138 kvadratmeter beliggende ved Bøjden, der er en mindre havneby på det sydvestlige Fyn. Boligerne er markeret med et kryds på kortet i Figur 5.1.

Figur 5.1

Estimationsgrundlag for en ejerlejlighed i Odense, et parcelhus i Nyborg og en landejendom ved Bøjden



Anm.: Til hver af de tre boliger er området indeholdende de 15.000 nærmeste bolighandler i perioden 1992-2021 indtegnet.  
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

For alle tre boliger skønnes boligens kvadratmeterpris ud fra de 15.000 nærmeste boliger, der er handlet i perioden 1992-2021, jf. Tabel 5.1. På kortet ovenfor er naboområdet med handlede boliger indtegnet.

For lejligheden i Odense ligger de solgte boliger inden for en afstand af 2,42 km. af lejligheden selv. Alle boligerne er beliggende i Odense kommune, og handlerne omfatter både ejerlejligheder og parcelhuse. Kvadratmeterprisen tillades dog at variere alene af den grund, at der er tale om forskellige boligtyper.

Datagrundlaget for parcelhuset i Nyborg indeholder boliger i op til 20,7 km. fra huset. Dette følger af, at Nyborg er en mindre by end Odense, og at tætheden for boliger i området omkring Nyborg er lavere. Området overlapper flere kommuner.

For ejendommen ved Bøjden findes de handlede boliger inden for 26,2 km. På halvøen Horne Land er tætheden af boliger forholdsvis lav, og boligprisen vil derfor i højere grad være fastsat ud fra nærliggende byer som landsbyen Horne og havnebyen Faaborg. Handlede boliger på blandt andre øerne Lyø og Helnæs indgår ligeledes i prisvurderingen. Også byer i lidt længere afstand indgår i prisfastsættelsen, blandt andre Haarby og Assens. Den geografiske vægtning bestemt ved afstanden til landejendommen i Bøjden sikrer, at der lægges størst vægt på lokale forhold.

Til hver af de tre boliger haves et estimationsgrundlag på 15.000 bolighandler. For hver handel haves kvadratmeterprisen og en række baggrundskarakteristika på boligen. De anvendte karakteristika er angivet i Tabel 5.4 bagerst i kapitlet.

Det er således muligt at sammenholde karakteristika på disse faktisk solgte boliger med karakteristika for hver af de tre betragtede ejendomme. Med andre ord tages der hensyn til, at de solgte boliger kan være anderledes end de betragtede boliger i forhold til for eksempel størrelse, opførelsesår og beliggenhed.

De handlede boliger i området omkring hver af de betragtede boliger opdeles tilfældigt i to. Det ene datasæt indeholder 12.000 boliger, hvorpå sammenhængen mellem den betragtede boligs karakteristika og kvadratmeterprisen fastsættes. Det andet datasæt indeholder de resterende 3.000 boliger. Dette datasæt anvendes undervejs i processen til at sikre, at den fundne sammenhæng mellem karakteristika og pris kan tolkes som generel og ikke i for høj grad tilpasses netop de betragtede handler.

Sammenhængen mellem boligens karakteristika og kvadratmeterprisen fastsættes ved en avanceret regressionsmodel. Der udføres én unik regression for hver enkelt ejerbolig. Regressionsmodellen ved hver bolig adskiller sig ved det anvendte datagrundlag.

Datagrundlaget for estimationer af to tætbeliggende ejendomme vil kun adskille sig ved få boliger i udkanten af området. Hvis for eksempel nabohuset til parcelhuset i Nyborg har tæt på samme karakteristika, vil den vurderede kvadratmeterpris ved disse to boliger kun afvige ganske lidt fra hinanden. Hvis to nært beliggende boliger i stedet har forskellige karakteristika, kan kvadratmeterprisen være forskellig. Det kan for eksempel være, at parcelhuset i Nyborg ligger med udsigt over Storebælt, mens nabohuset ligger i anden række og derfor ingen udsigt har. Herudover er husene ens. Da vil den betragtede bolig have en værdi, der er forskellig fra, men stadig relativt tæt på naboboligens pris. Den ændrede pris for boligen i første række til havet afspejler dermed værdiforskellen mellem et hus med havudsigt og et uden. Således bliver hvert karakteristika for den aktuelle bolig sammenlignet med kendetegnene for de solgte boliger.

Træfsikkerheden af den estimerede model bestemmes ud fra to mål: RMSE og PM20 (defineres senere i kapitlet).

PM20 er for lejligheden i Odense på 84 pct. Det vil sige, at når modellen testes på de 3.000 boliger, som ikke anvendes i selve modellen, så ligger den estimerede kvadratmeterpris indenfor +/- 20 pct. af den faktisk handlede pris i 84 pct. af tilfældene. Jo tættere PM20 er på 100, jo bedre. For huset i Nyborg og ved Bøjden er PM20 henholdsvis 69 og 62 pct.

For lejligheden i Odense er RMSE 0,18, mens den for de øvrige boliger er 0,27 og 0,30 for henholdsvis Nyborg og Bøjden. RMSE skal være så tæt ved nul som muligt.

Disse mål for træfsikkerhed indikerer, at prisen på lejligheden i Odense er bestemt mest sikkert. Generelt gælder det, at jo mindre afstand, der er til de nærmeste solgte boliger, desto mere præcis bliver den estimerede pris. Det betyder, at træfsikkerheden i større byer generelt er højere end i de mindre byer og på landet.

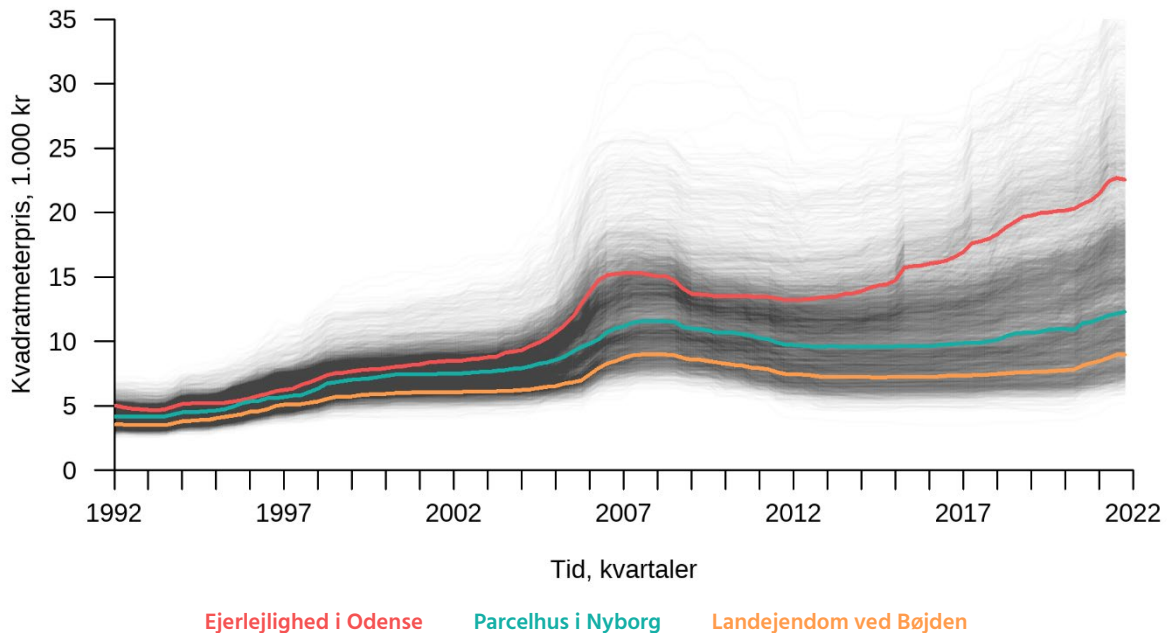
Er boligen handlet i løbet af den betragtede periode, kan afvigelsen mellem den faktiske og den vurderede handelspris ligeledes beregnes. Lejligheden i Odense blev i tredje kvartal 2012 solgt til en kvadratmeterpris på 13.420 kr. Den af modellen prædikterede kvadratmeterpris i dette kvartal er 13.010 kr. Det vil sige, at boligens estimerede pris bliver undervurderet med cirka 3 pct.

Parcelhuset i Nyborg blev i fjerde kvartal 2019 solgt til en kvadratmeterpris på 11.560 kr., mens den prædikterede kvadratmeterpris i samme kvartal er 9.960 kr. Denne bolig bliver således undervurderet med 13,8 pct. Landejendommen i Bøjden overvurderes med 16,1 pct. i forhold til den faktiske kvadratmeterpris på 2.390 kr. i andet kvartal 1995 (estimeret: 2.780 kr./m<sup>2</sup>).

Ved estimationen vurderes boligens kvadratmeterpris i alle kvartaler i løbet af den betragtede periode, uanset om boligen er handlet eller ej. Figur 5.2 viser den estimerede prisudvikling for en ejerlejlighed i Odense, et parcelhus i Nyborg, en landejendom på Sydvestfyn samt yderligere 5.000 tilfældigt udvalgte ejerboliger på Fyn (den grå sky). Boligprisstigningen i årene frem mod 2007 og i perioden fra 2013 vurderes kraftigst i Odense.

**Figur 5.2**

**Kvadratmeterpris for en ejerlejlighed i Odense, et parcelhus i Nyborg, en landejendom ved Bøjden og 5.000 tilfældigt udvalgte ejerboliger på Fyn**



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

## 5.2 Geografisk vægtet regression

Den benyttede metode til skøn over prisen på ejerboliger er inspireret af geografisk vægtet regression (Geographically Weighted Regression, GWR), jf. Brunsdon, Fotheringham, & Charlton (1966).

Ideen i geografisk vægtet regression er at anvende en boligs geografiske placering, så analysen udelukkende baseres på ejendomme i naboområdet til boligen selv. Ydermere tildeles hver bolig i naboområdet en vægt, der afhænger af afstanden til den betragtede bolig. Her ved vil metoden give en god beskrivelse af lokale forhold, som kan være medforklarende til variation i prisfastsættelsen af en bolig.

Ved vurdering af værdien af en ejerbolig er den traditionelle tilgang at foretage en lineær regression, hvor kvadratmeterprisen på boligen søges forklaret ud fra en række kendetegn ved boligen og forskellige indikatorer for boligens beliggenhed. Boligkarakteristika kan for eksempel omfatte boligens størrelse, opførelsesår, grundstørrelse mv., mens geografiske karakteristika kan være afstand til kyst, motorvej m.fl. Analysen foretages for et afgrænset geografisk område, hvilket kan være én samlet regression for hele landet eller flere regressioner

udført indenfor eksempelvis hver af landets fem regioner. Sådanne vurderingsmetoder er kendt som hedoniske prismodeller, jf. Court (1939) og Rothenberg m.fl. (1991).

I modsætning til traditionelle prismodeller inkluderes i denne undersøgelse også beboerkaraktistika, som blandt andet omfatter husstandsindkomst, og antal år de sælgende ejere har ejet boligen op til bolighandlen. Beboerkaraktistika medtages som forklarende variable for at afsløre boligens standsmæssige egenskaber, der ellers ikke er observerbar i data. Der formodes således en sammenhæng mellem ejernes husstandsindkomst og boligens vedligeholdelsesmæssige stand.

Estimeres sammenhængen mellem ejerboligers handelspris og baggrundsvariablene ved én lineær regression, som indeholder bolighandler fra et større geografisk område, haves en såkaldt global model. Herved antages implicit, at den marginale priseffekt af de indeholdte boligkaraktistika er den samme i hele området. Hvis man for eksempel foretager en analyse for hele Danmark, da ville man implicit antage, at den marginale effekt af at være en kystnær bolig er den samme ved Strandvejen i Hellerup som på Ærø i det Sydfynske Øhav. Ved en global model vil den marginale effekt af at ligge tæt ved en motorvej ligeledes være den samme for en bolig, der ligger 300 m. fra motorvejen i et åbent område i Jylland (og derfor er støjplaget) og en bolig, som ligger tilsvarende tæt ved en motorvej i Københavnsområdet (men måske i fjerde husrække fra motorvejen, hvorfor boligen kun i begrænset omfang er udsat for støj).

Det virker således ikke hensigtsmæssigt, at priseffekterne af enkelte boligkaraktistika er ens for større geografiske områder. Det vi ønsker er derfor en lokal model, hvor der tages hensyn til boligmarkedets betydelige geografiske forskelle på tværs af landet.

Geografisk vægtet regression er en lokal model. Ideen er, at der udføres én regression for hver enkelt bolig, hvor værdien ønskes fastlagt. Estimationens datagrundlag er handlede ejerboliger i et relativt lille nærområde til den betragtede bolig. Herved bestemmes sammenhængen mellem ejerboligers handelspris og baggrundsvariablene i netop det område, hvor boligen er beliggende. Der opnås således en meget fin beskrivelse af de observerede lokale forhold.

### Ejendomsværdi ved geografisk vægtet regression

Betragt en ejerbolig  $i$ , hvor ejendomsværdien ønskes vurderet<sup>10</sup>. Værdien ønskes fastsat i hver tidsenhed inden for en given periode. Tidsenheder angives ved  $t \in \{0, 1, \dots, \tau\}$ . Boligen er ikke nødvendigvis handlet inden for perioden, men det kan være tilfældet (og boligen kan være handlet mere end én gang).

Metoden beskrevet nedenfor udføres selvstændigt for hver enkelt bolig  $i \in \{1, 2, \dots, I\}$ , hvor boligprisen ønskes vurderet. Metoden er således ganske beregningstung.

Det antages, at (den naturlige logaritme til) en boligs kvadratmeterpris til tidspunkt  $t$ ,  $p_{i,t}$ , er en lineær funktion af en række baggrundskaraktistika på boligen,  $X_i$ , og en konstant, som bestemmer prisen i den angivne tidsperiode:

$$(5.1) \quad p_{i,t} = \beta_i X_i + \alpha_{i,t}$$

hvor  $\beta_i$  er en vektor bestående af koefficienter hørende til hvert karakteristika,  $\alpha_{i,t}$  er en vektor bestående af koefficienter hørende til hver tidsenhed og  $t \in \{0, 1, \dots, \tau\}$ .

---

<sup>10</sup> Beskrivelsen i dette afsnit tager udgangspunkt i, at regressionen udføres som en lineær regression (ordinary least squares, OLS). Dette er tilfældet i tidligere undersøgelser udført af DREAM, jf. Hansen, Iversen og Stephensen (2018) og Stephensen, Hansen, Jørgensen og Hansen (2015). Regressionen kan dog udføres med andre regressionsmetoder som beskrevet i afsnit 5.3.

I ligning (5.1) bestemmer  $\beta_i X_i$  niveauet for boligens handelspris. Niveauet for handelsprisen kan for eksempel være relativt højt, hvis boligen har en attraktiv beliggenhed, eller prisen kan være forholdsvis lav, hvis boligen er opført i dårlige materialer. Boligens prisudvikling over tid bestemmes af  $\alpha_{i,t}$ .

Idet bolig  $i$  ikke nødvendigvis er handlet indenfor den betragtede periode, er det ikke sikkert, at handelsprisen  $p_{i,t}$  kan identificeres. Regressionsanalysen angivet ved ligning (4.1) udføres derfor på et datagrundlag omfattende handlede ejerboliger i nærområdet til bolig  $i$ . Estimationen sker på de i alt  $N_i$  bolighandler, som ligger tættest ved bolig  $i$ . Denne mængde af boliger betegnes  $\mathbb{N}_i$ , og en bolig  $j$  i denne mængde betegnes  $j \in \mathbb{N}_i$ . Såfremt bolig  $i$  er handlet indenfor den betragtede periode, vil boligen selv være indeholdt i  $\mathbb{N}_i$  (men det behøver ikke være tilfældet).

Fastsættelse af størrelsen af nærområdet med handlede boliger beskrives i afsnit 5.4. Som det fremgår heraf, indgår der i praksis 12.000 handler i hver enkelt estimation. Herved bliver det nærområde, som estimationen foretages på baggrund af, typisk relativt stort. Ofte har området en radius på over 10 km. De forholdsvis store områder sikrer, at der er tilstrækkeligt data til at estimere modellen. Der inkluderes en vægtning af hver bolighandel således, at handler tæt ved bolig  $i$  vægtes højt, hvorfra vægtningen af den enkelte handel aftager, når afstanden til den betragtede bolig øges. Dette sikrer modellens gode lokale egenskaber.

Ejendomsværdien for bolig  $i$  fastlægges herefter i to skridt: Først bestemmes den lokale sammenhæng mellem prisen på de handlede ejerboliger i  $\mathbb{N}_i$  og naboboligernes baggrundskarakteristika. Dernæst anvendes den estimerede sammenhæng til at imputere handelsprisen på den betragtede ejerbolig.

I første skridt antages det, at der er en lineær sammenhæng mellem (den naturlige logaritme til) en boligs kvadratmeterpris og baggrundskarakteristika inklusiv indikatorvariable for handelstidspunkt:

$$(5.2) \quad p_j = \beta_j X_j + \alpha_j T_j + \epsilon_j \quad \text{for alle handlede boliger } j \in \mathbb{N}_i$$

De anvendte baggrundsvARIABLE i  $X_j$  omfatter et konstantled, geografiske karakteristika, boligkarakteristika og beboerkarakteristika. De anvendte karakteristika kan ses i Tabel 5.4 bagerst i dette kapitel.

$T_j$  indeholder  $\tau - 1$  indikatorvariable. Indikatorvariablen hørende til den tidsenhed, hvor bolighandler er fundet sted, er lig én, mens øvrige indikatorvariable er lig nul (så variablene i  $T_j$  er såkaldte tidsdummies). Bemærk, at da de handlede ejerboliger i  $\mathbb{N}_i$  er udvalgt individuelt til regressionen omfattende bolig  $i$ , da vil estimationsgrundlaget – og dermed de estimerede parameterestimer – være unikke for hver enkelt bolig.

I første trin estimeres ligning (5.2) for bolig  $i$  på baggrund af de  $N_i$  solgte ejerboliger i  $\mathbb{N}_i$ . Der udføres en vægtet regression, hvor  $w = (w_1, w_2, \dots, w_{N_i})$  indeholder en vægt af hver enkelt bolighandel. Vægten  $w_j$  bestemmes ud fra den geografiske afstand mellem den betragtede ejerbolig  $i$  og hver af de handlede boliger  $j \in \mathbb{N}_i$ . Vægten er størst for boliger tæt ved den betragtede bolig og aftager jo større afstand, der er mellem den handlede bolig  $j$  og boligen, hvis pris ønskes fastsat,  $i$ . Fastsættelse af vægtningen beskrives nærmere i afsnit 5.4. Estimationen foretages på et datagrundlag, som er unikt for den  $i$ 'te bolig, hvorfor resultatet er et sæt af parameterestimer,  $(\hat{\beta}_i, \hat{\alpha}_{i,t})$ , som er entydigt for den betragtede bolig  $i$ . En simpel global model ville derimod finde fælles parameterestimer for alle ejerboliger.

I andet trin imputeres kvadratmeterprisen for den betragtede bolig  $i$  ud fra modellen (5.2) estimeret på de  $N_i$  handlede boliger tættest ved boligen selv. Kvadratmeterprisen for bolig  $i$  bestemmes i periode  $t$  til:



$$(5.3) \quad p_{i,t} = \hat{\beta}_i X_i + \hat{\alpha}_{i,t} \quad \text{for alle tidsperioder } t \in \{0, 1, \dots, \tau\}.$$

Ligning (5.3) fastsætter boligens værdi på et givet tidspunkt  $t$ . Ved at gennemløbe alle tidsenheder opnås et bud på udviklingen i boligens kvadratmeterpris gennem den analyserede periode.

Den anvendte implementering af geografisk vægtet regression er udviklet i DREAM og programmeret i statistikprogrammet R. Implementeringen er baseret på nabosøgningsalgoritmen ANN (Approximate Nearest Neighbor, der dog også effektivt kan finde eksakte naboer), som er en hurtig metode til at finde den enkelte boligs nærmeste naboer.

## 5.3 Sammenligning af modeller og valg af modeltype

DREAM har tidligere vurderet prisudviklingen på ejerboliger ved geografisk vægtet regression, jf. Hansen, Iversen & Stephensen (2018). Her blev regressionen udført ved lineær regression som beskrevet i afsnit 5.2.

Internationalt er det de seneste år forsøgt at forbedre metoden. Den geografiske vægtning i forhold til et center anvendes fortsat. Men den lineære regressionsmodel er fundet for simpel, hvorfor den er forsøgt erstattet med moderne teknikker indeholdende maskinlæring. Disse modeller kaldes *Geographically Weighted Machine Learning* og anvendes blandt andet af Chen m.fl. (2018), Li (2019), Quiñones m.fl. (2021) og Xu m.fl. (2021).

Metoder baseret på beslutningstræer har vist sig effektive og er blandt de mest populære valg til analyse af geografiske problemstillinger. Træbaserede algoritmer kan anvendes til regressionsanalyser ved at opdele i grupper efter baggrundskarakteristika. For hver gruppe fastsættes værdien af responsvariablen, som i nærværende tilfælde er boligens kvadratmeterpris.

Figur 5.3 viser et eksempel på et simpelt beslutningstræ, hvor kvadratmeterprisen for en ejerbolig bestemmes ud fra tre baggrundskarakteristika: Kommune som 'storby' eller 'øvrige', boligens opførelsesår og boligareal. Træet dannes ud fra den øverste startnode, som indeholder samtlige 58.000 handlede ejerboliger i 2020. Den gennemsnitlige handelspris er knap 21.500 kr. pr. kvadratmeter.

I hvert dybdeniveau opdeles træet i to grene<sup>11</sup>. Der splittes efter det baggrundskarakteristika, som fører til den maksimale reduktion af tabsfunktionen, hvilket også kaldes den største mulige formindskelse af urenheden (eng: impurity). I princippet undersøges alle mulige split af samtlige variable. Da dette kan være en tidskrævende proces anvender de fleste implementeringer approksimative algoritmer, som ikke nødvendigvis finder optimale split.

I eksemplet i figuren nedenfor opdeles først efter, om boligen er beliggende i en storbykommune eller ej. Kvadratmeterprisen i storbyerne ses at være dobbelt så stor som i de øvrige kommuner.

For boliger i storbykommunerne ses yderligere split efter først opførelsesår, idet der opdeles efter boliger opført før 1939 og øvrige. Tidligt opførte boliger har den højeste kvadratmeterpris. Dette kan skyldes, at boliger opført før 1940 typisk vil være attraktive beliggende tæt ved bycentrum. For boliger opført efter 1939 sker yderligere et split efter opførelsesår, idet der skelnes mellem, om boligen er opført før eller efter 2004. Ejerboliger opført i perioden 1940-2003 opdeles efter boligareal. Den mindste kategori med et boligareal på højst 94 m<sup>2</sup>

---

<sup>11</sup> I rapporten betragtes udelukkende binære beslutningstræer, dvs. hvert split indeholder præcis to grene.

ses at have den højeste pris. Dette kan tolkes som ejerlejligheder, som typisk handles til en højere kvadratmeterpris end parcelhuse.

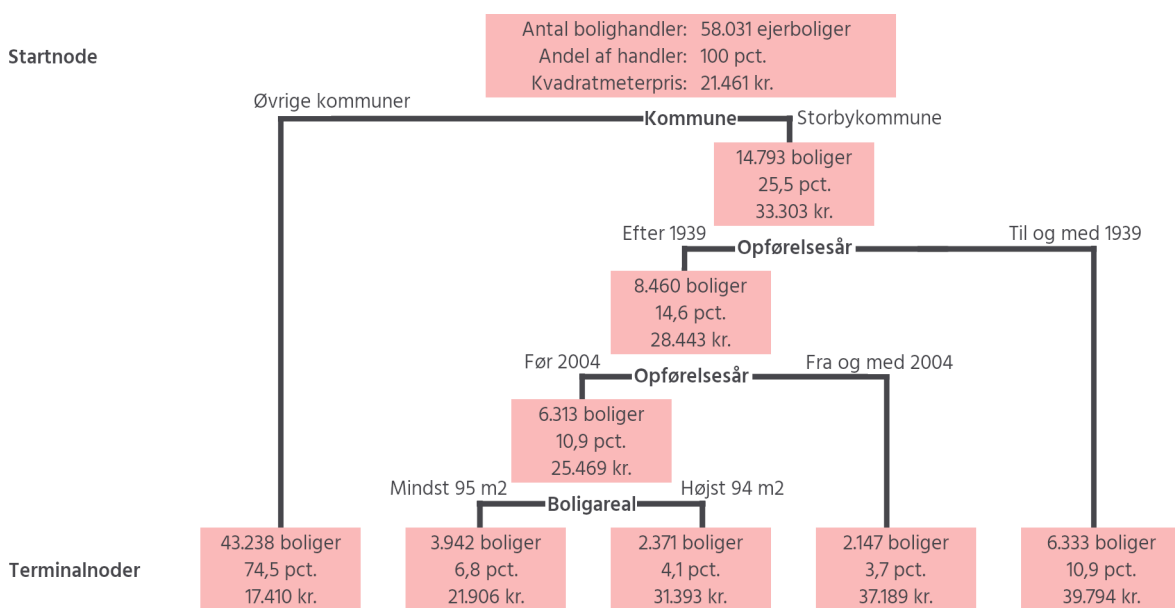
Træer stopper med at gro, når et stopkriterie er opfyldt. Dette kan være, når et fast antal split (angivet ved parameteren  $d$ ) er foretaget<sup>12</sup>. Da hvert split danner præcis to nye grene, kan dette alternativt formuleres som, at træet stopper med at gro, når et givet antal terminalnoder er nået (når antal terminalnoder er lig  $d + 1$ ).

De sidste terminalnoder, hvor der ikke opdeles yderligere, indeholder forudsigelsen af i vores tilfælde kvadratmeterprisen for en bolig med de givne karakteristika. I eksemplet foretages fire split, så træet indeholder fem terminalnoder.

En fordel ved beslutningstræer er, at de kan konstrueres med både kontinuerte og kategoriske variable. Desuden tillader modellerne en ikke-lineær sammenhæng mellem responsvariablen (boligpris) og baggrundsvARIABLE (boligkarakteristika), ligesom beslutningstræer potentielt indeholder fuld interaktion mellem alle variable.

Metoder som blot estimerer et enkelt beslutningstræ har vist sig mindre effektive. Derimod har metoder, som baserer sig på mange træer vist sig at mindske prædiktionsfejlen. Boks 5.1 indeholder en mere teknisk beskrivelse af træbaserede algoritmer.

**Figur 5.3**  
**Eksempel på beslutningstræ**



Anm.: Figuren viser et simpelt beslutningstræ, som prædikerer ejerboligers kvadratmeterpris ud fra bolighandler i 2020. Træet kan opdele efter tre baggrundsvARIABLE: kategorisk kommunevariabel med udfaldsrum {storby, øvrige kommuner} (storby er København, Frederiksberg, Aarhus, Odense og Aalborg kommuner) samt to kontinuerte variable (heltal) ved boligens opførelsesår og boligareal i kvadratmeter.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

<sup>12</sup> Der anvendes typisk et stopkriterie for at undgå meget store træer med mange terminalnoder. Store træer kan medføre overfitting. Andre stopkriterier end antal split/antal terminalnoder kan anvendes. Øvrige stopkriterier er typisk en maksimal dybde eller et minimum antal observationer i hver terminalnode. Det er også muligt at lade træet gro til fuld størrelse for derefter at beskære det (eng: pruning), denne tilgang beskrives dog ikke nærmere i nærværende rapport.

## Boks 5.1

### Træbaserede algoritmer

Algoritmer, der blot estimerer ét beslutningstræ, har ikke vist sig lige så effektive som metoder, der kombinerer mange træer.

I algoritmer med mange beslutningstræer vil hvert træ ofte være en svag klassifikationsmodel, som isoleret set klarer sig dårligt (på engelsk kaldet en *weak learner*). Men kombineret giver disse træer en model, der sædvanligvis opnår god prædiktionssevne. Ofte anvendes algoritmer, som enten er baseret på såkaldt *bagging* eller *boosting*. Begge metoder kan anvendes på både regressions- og klassifikationsproblemer, dvs. til henholdsvis at vurdere størrelsen af et tal og skønne en type ud fra flere mulige udfald.

Ved *bagging* tages gennemsnittet af mange beslutningstræer. Hvert træ estimeres på et unikt datasæt af samme størrelse som det oprindelige datasæt. Estimationsdata dannes ved tilfældigt at trække (med tilbagelægning) fra det oprindelige datasæt (kaldet *bootstrap*). Nogle observationer vil optræde flere gange, mens andre observationer slet ikke vil indgå i estimationsdata. Træets prædiktionssevne kan testes på de observationer, som ikke er udtrukket til estimationsdata (kaldet *out-of-bag examples*). Kun et udvalg af det totale antal forklarende variable indgår i det enkelte træ. Et eksempel på en algoritme, som anvender *bagging*, er *random forest*.

Ideen i *boosting* er, at modellen løbende forsøger at mindske fejlen i forhold til den foregående model. Beslutningstræerne estimeres her sekventielt. Det første træ prædikterer responsvariabelen. Det andet træ anvender forskellen mellem forudsigelserne fra det første træ og sandheden (residualerne) som responsvariabel. Herved kan modellen forbedres i de områder, hvor den ikke klarer sig godt. Det tredje træ estimeres på residualerne fra det andet træ og så fremdeles. Herved vil algoritmen løbende forbedres. *Boosting* teknikken beskrives ligeledes i afsnit 5.4.

Som alternativ til lineær regression testes tre træbaserede algoritmer, som alle baseres på *boosting*. Metoderne adskiller sig ved implementeringen.

*Extreme Gradient Boosting* (XGBoost) er en meget anvendt implementering, som forsøger at modvirke overfitting ved at straffe komplicerede modeller. En ulempe ved XGBoost er, at den ikke kan håndtere kategoriske variable. *Light Gradient Boosting Machine* (LightGBM) anvender en anden metode til at gro træer (leaf-wise også kaldet *best-first*), hvilket giver en kortere afviklingstid typisk mod lidt større prædiktionsfejl. LightGBM kan håndtere kategoriske variable, hvor *CatBoost* dog på det seneste har skubbet rammen for at håndtere kategoriske variable.

Kilde: Egen tilblivelse på baggrund af James, Witten, Hastie & Tibshirani (2021), Chen m.fl. (2016), Ke m.fl. (2017) og Prokhorenkova (2018).

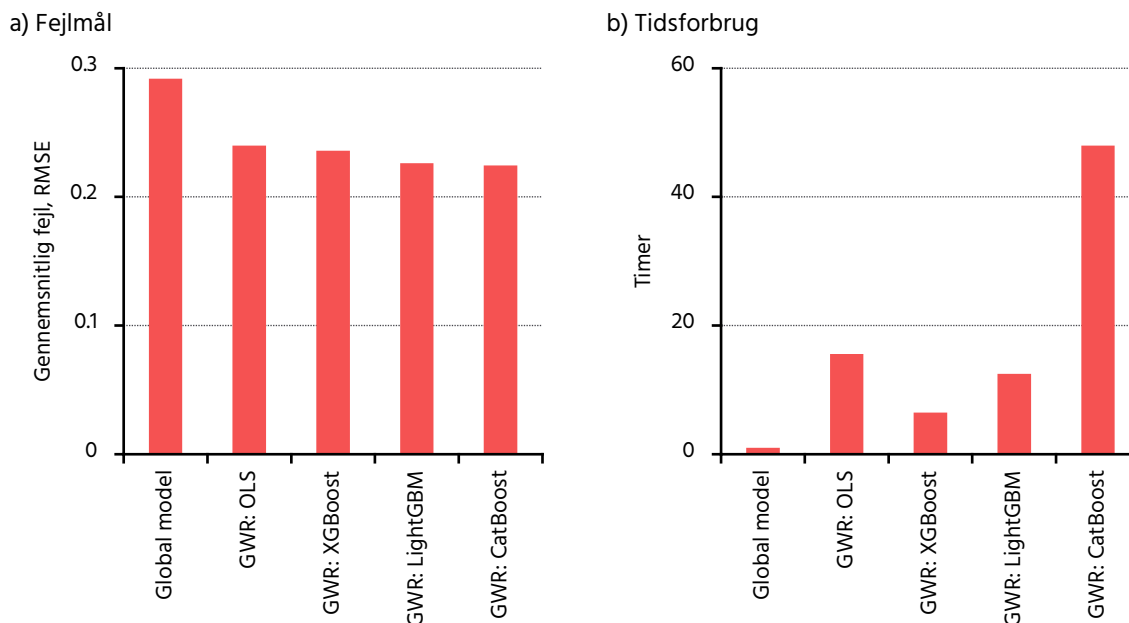
Der er udført en sammenligning af forskellige modeller for fastsættelse af kvadratmeterprisen for en ejerbolig, jf. Figur 5.4. Sammenligningen er udført ved at skønne kvadratmeterprisen for 20.000 tilfældigt udvalgte boliger, som er handlet i perioden 1999-2015. Den vurderede kvadratmeterpris sammenlignes herefter med den faktiske handelspris, og figuren viser den gennemsnitlige afvigelse mellem disse vurderet ved root-mean-square error (RMSE). Den bedste model er den, som giver den laveste fejl.

Den simpleste model er en global model, som estimerer en lineær regressionsmodel for hele Danmark. Denne model opnår den største gennemsnitlige afvigelse, men kørselstiden er hurtig. Herefter følger geografisk vægtet lineær regression, som tidligere er anvendt til at fastlægge boligprisudviklingen.

Herudover er udført tre modelvarianter, som ligeledes anvender geografisk vægtning, men hvor regressionen udføres med træbaserede modeller, jf. Boks 5.1. Det ses af Figur 5.4, at metoderne LightGBM og CatBoost klarer sig nærmest lige godt, dog giver CatBoost en marginalt lavere gennemsnitlig fejl. Tidsforbruget ved CatBoost er imidlertid næsten fire gange større end ved LightGBM, hvorfor sidstnævnte metode vælges til vurdering af boligprisudviklingen, hvor et stort antal estimationer gennemføres.

Figur 5.4

#### Sammenligning af udvalgte statistiske modeller



Anm.: Modellerne er estimeret på 20.000 boliger handlet i perioden 1999-2015. Ved en global model estimeres én lineær regression på samtlige cirka 850.000 bolighandler i perioden. Ved GWR-modellerne foretages en estimation for hver af de 20.000 boliger, og estimationen foretages på de 10.000 nærmeste handlede boliger, som vægtes med afstanden til den betragtede bolig. De fire GWR-modeller adskiller sig udelukkende i forhold til, hvilken regressionsmetode, som anvendes.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

## 5.4 Tuning af modellen

Ud fra en afvejning af træfsikkerhed og kørselstid anvendes estimationsmetoden LightGBM. Det er nødvendigt at fastsætte, hvor stort et naboopråde, som skal indgå i den enkelte estimation. Desuden kræver algoritmen, at to såkaldte hyperparametre bestemmes. Disse parametre styrer metodens indlæringsproces.

### Naboområdets størrelse

Geografisk vægtning indebærer, at der for alle betragtede boliger skal findes en mængde af nærmeste boligsalg, hvorpå regressionsmodellen estimeres. Til at danne naboområdet med handlede boliger anvendes ANN-algoritmen (Approximate Nearest Neighbor), jf. Arya m.fl. (1998) og (2023). På trods af navnet er metoden også effektiv til at finde eksakte nærmeste naboer.

Mængden af nabosalg til hver betragtet bolig bestemmes ud fra fire parametre:

1. Det mindste antal accepterede salg  $k^{min}$ ,
2. Et maksimalt antal salg  $k^{max}$ ,
3. En søgeradius, som handlede boliger findes indenfor,  $r$ , og
4. Den geografiske vægtning af de enkelte handler (kernefunktionens form).

Den geografiske vægtning af de enkelte handler i estimationsgrundlaget foretages ved kernefunktionen kaldet bisquare<sup>13</sup>. Ved afstand nul giver funktionen den maksimale vægt på én. Omkring dette centrum vægter bisquare-funktionen stort set på samme måde som den gaussiske kernefunktion<sup>14</sup>. Fordelen ved bisquare er, at vægten er nul ved fuld radiusafstand, hvilket er i modsætning til den gaussiske kerne.

De tre øvrige parametre indeholdt i estimationerne optimeres ved grid search. Her findes et godt sæt af parameterværdier ved at foretage en række estimationer med forskellige kombinationer af parameterværdier. Herudfra udvælges kombinationen, som giver den bedste træfsikkerhed.

Træfsikkerheden ved forskellige parametervalg sammenlignes ved at inddele landet efter otte bystørrelser og Bornholm. For hver gruppe køres en serie estimationer på 2.500 tilfældigt udvalgte boliger, som blev handlet inden for den betragtede periode. I en sådan serie varieres værdien af de tre parametre inden for givne intervalgrænser i alle tænkelige kombinationer.

**Tabel 5.1**  
 Valg af parameterværdier som fastsætter naboområdet størrelse

Område	Antal handler 1992-2021	Antal estimation	Antal test prædiktio	Radius i km.	RMSE	PM20 i pct.	MAPE pct.
<b>Efter bystørrelse ved antal indbyggere (eksklusiv Bornholm)</b>							
Hovedstadsomr.	292.905	12.000	3.000	2,7	0,17	86,1	12,0
100.000 eller flere	138.389	12.000	3.000	3,5	0,17	85,4	12,1
50.000-99.999	105.590	12.000	3.000	6,1	0,20	81,1	14,1
20.000-49.999	194.950	12.000	3.000	10,3	0,19	80,5	13,7
10.000-19.999	105.468	12.000	3.000	14,5	0,21	78,6	14,9
5.000-9.999	107.511	12.000	3.000	16,8	0,22	77,1	15,8
1.000-4.999	227.702	12.000	3.000	15,4	0,24	73,3	17,6
Færre end 1.000	247.196	12.000	3.000	17,2	0,35	55,6	27,2
<b>Bornholm</b>							
Hele øen	14.537	11.630	2.907	28,7	0,31	57,6	24,1

Anm.: Bornholm køres som et selvstændigt område, mens den resterende del af landet er opdelt efter antal indbyggere i det bymæssige område, hvor boligen er beliggende.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

<sup>13</sup> Bisquare-kernefunktionen har den funktionelle form  $f(x) = \max\left(1 - \left(\frac{x}{c}\right)^2, 0\right)$ , hvor  $c$  angiver, hvorfra funktionsværdien antager værdien nul (søgeradius). I vores anvendelse sættes  $c$  lig afstanden til den yderste bolig i estimationsgrundlaget.

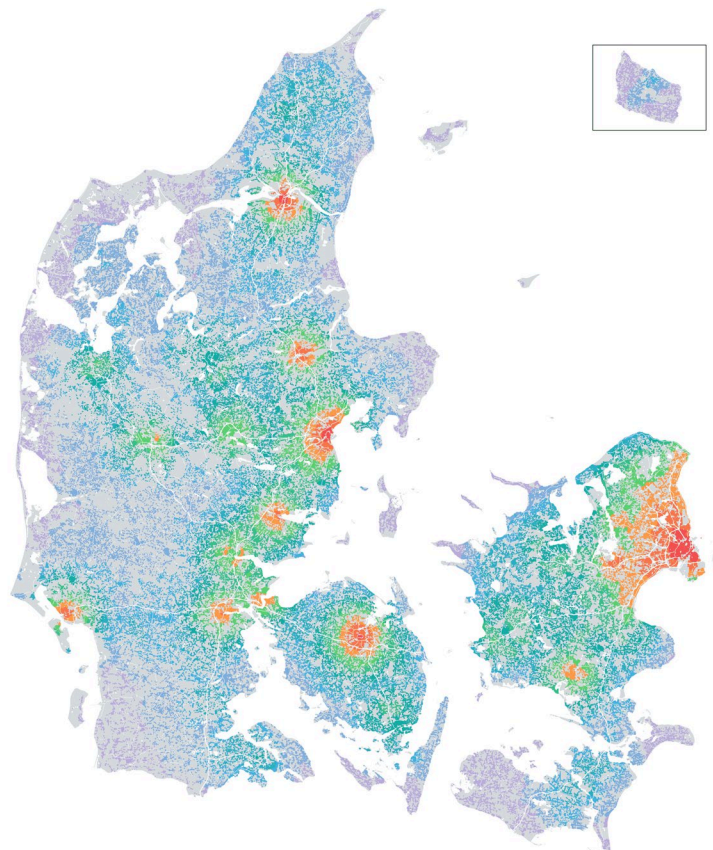
<sup>14</sup> Den Gaussiske kernefunktion har generelt den funktionelle form  $f(x) = a \exp\left(-\frac{(x-b)^2}{2c^2}\right)$ , hvor  $a$  bestemmer højden af kurvens toppunkt,  $b$  fastlægger  $x$ -værdien for klokkes toppunkt og  $c$  afgør klokkes bredde (søgeradius). I geografisk regression sættes  $a = 1$  og  $b = 0$ , mens bredden  $c$  kan optimeres ud fra den givne problemstilling.

Ved geografisk vægtet regression estimeres en handelspris for hver kombination af parameterværdier, idet boligens faktiske handelspris forudsættes at være ukendt i selve estimationsprocessen. Estimationsresultatet ved hvert sæt af parameterværdier kan da sammenlignes med boligens faktiske handelspris.

Træfsikkerheden måles ved root-mean-square error (RMSE). Som supplement hertil angives to mere let fortolkelige mål i form af den såkaldte PM20 og den gennemsnitlige absolutte procentvise afvigelse (mean absolute percentage error, MAPE). PM20 angiver, hvor stor en andel af de prædikterede kvadratmeterpriser, som ligger indenfor +/- 20 pct. af den faktiske handelspris. Træfsikkerheden er størst, når RMSE og MAPE er så lille som mulig, og PM20 er størst mulig. Disse krav er ikke nødvendigvis opfyldt for det samme sæt tuningsparameter-værdier.

Ved fastsættelse af størrelsen af naboområdet med handlede boliger er det valgt at anvende samme sæt af parameterværdier for alle øvrige dele af landet end Bornholm. Her anvendes et naboområde bestående af de 15.000 nærmeste handlede boliger. Heraf indgår de 80 pct. i selve estimationen (træningsdata), mens de resterende 20 pct. anvendes til at sikre at modellens resultater kan generaliseres (testdata).

**Figur 5.5**  
Søgeradius



Op til 2½ km. 2½-5 km. 5-7½ km. 7½-10 km. 10-12½ km. 12½-15 km. 15-17½ km. 17½-20 km. 20-25 km. Over 25 km.

Anm.: Figuren viser afstanden i kilometer til den yderste bolig i estimationsgrundlaget.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

At anvende samme parameterverdier for hele landet følger af, at et naboopråde bestående af 15.000 handlede boliger generelt giver god træfsikkerhed. I de tilfælde, hvor dette valg ikke findes optimalt, øges prædiktionsfejlen kun ganske svagt. Derfor vælges samme parameterverdier for hele landet undtagen Bornholm.

På Bornholm er det fundet mest hensigtsmæssigt at foretage én samlet estimation, hvor alle cirka 14.500 handler i den betragtede periode anvendes. Heraf indgår 80 pct. i træningsdata og 20 pct. i testdata.

Den bedste træfsikkerhed opnås i Hovedstadsområdet samt Aarhus, Odense og Aalborg. Det fremgår af de tre kolonner yderst til højre i Tabel 5.1. Modellens prædiktionssevne bliver svagere, når mindre byområder betragtes. Dette skyldes større afstand til handlede boliger og ofte også større variation mellem boligerne, end det er tilfældet i tættere bebyggede områder.

Søgeradius følger bystørrelsen og den naturlige tæthed af ejerboliger, jf. Figur 5.5

### Fastsættelse af hyperparametre

Frem for at estimere et stort beslutningstræ er ideen i såkaldte boosting algoritmer, at der i stedet foretages en langsom læringsproces, hvor modellen løbende optimeres. Optimeringen sker ved, at en række beslutningstræer bestemmes sekventielt, idet hvert træ estimeres ud fra information fra tidligere estimerede træer. Boosting algoritmer vil således kombinere et stort antal beslutningstræer.

Ved anvendelse af LightGBM er det nødvendigt at fastsætte værdien af flere såkaldte hyperparametre, dvs. parametre hvis værdi bruges til at styre metodens indlæringsproces. De vigtigste hyperparametre ved anvendelse af boosting algoritmer er antallet af beslutningstræer og algoritmens læringsrate<sup>15</sup>, jf. James, Witten, Hastie, & Tibshirani (2021). Boosting algoritmen og anvendelsen af de to hyperparametre skitseres i Boks 5.2.

Den optimale værdi af hyperparametrene kan ikke estimeres af modellen. Parameterverdierne findes i stedet ved tuning af modellen, dvs. der eksperimenteres med flere forskellige sæt af parameterverdier for at finde den kombination, som maksimerer modellens prædiktionssevne. Tuning af en model kaldes også hyperparameteroptimering.

Antallet af træer angiver antallet af iterationer i boosting algoritmen. Jo flere træer, som indgår, jo bedre prædiktionssevne kan modellen opnå. Ulempen ved at anvende mange træer kan være, at afviklingstiden øges, ligesom muligheden for såkaldt overfitting opstår. Overfitting er, at modellen i for høj grad tilpasses tilfældige udsving i data, den estimeres på, hvorfor modellen ikke nødvendigvis vil klare sig godt, når den anvendes på fremmed data. Det vil ikke være hensigtsmæssigt at anvende en sådan model til at skønne over boligprisudviklingen.

Generelt gælder det, at en langsom læringsproces giver den bedste prædiktionssevne. Læringsraten bestemmer, hvor hurtigt boosting algoritmen adopterer ny viden fra de træer, som indgår. Hvis læringsraten er lav, vil et stort antal træer være nødvendigt for at opnå en god prædiktionssevne. Og modsat vil et relativt lavt antal træer nødvendiggøre en høj læringsrate.

---

<sup>15</sup> Hvor store træer som indgår i algoritmen bestemmes ved antal split i hvert beslutningstræ,  $d$ . Dette kan ligeledes ses som en hyperparameter. Det antages, at  $d$  er lig 30. Idet beslutningstræerne er binære, vil hvert træ derved have  $d + 1 = 31$  terminalnoder.

## Boks 5.2

### Skitsering af metoden gradient boosting

Der er observeret en responsvariabel  $y = (y_1, y_2, \dots, y_n)$  og en række forklarende variable  $X = (X_1, X_2, \dots, X_n)$ . Det ønskes at bestemme sammenhængen mellem disse,  $y = f(X)$ . Ved gradient boosting udvikles en model,  $\hat{f}$ , ved sekventielt at konstruere beslutningstræer, som løbende tilføjes og forbedrer modellen.

Først dannes en simpel grundmodel. Herefter beregnes fejlen mellem grundmodellens prædiction og de observerede værdier dvs. residualerne.

Der dannes et beslutningstræ, som prædikerer residualerne. Træet vil typisk være relativt simpelt og indeholder få terminalnoder. Antal terminalnoder angives af brugeren ved parameteren  $d$ . Modellen opdateres med en nedskaleret version af dette træ. Herved forbedres modellen i de områder, hvor den prædikerer dårligt. At nedskalere med læringsraten  $0 < \lambda < 1$  sløver processen, hvilket hæmmer overfitting.

Modellen ajourføres således ved iterativt at bygge og tilføje træer, der forsøger at mindske fejlen i forhold til den foregående model. Dette fortsætter indtil antallet af træer når et brugervalgt maksimum,  $B$ . Alternativt kan metoden afbrydes, hvis prædiktionsfejlen ikke længere mindskes.

Algoritmen kan skitseres således:

1. Initialiser modellen. Typisk vælges  $\hat{f}^0(X) = 0$  eller  $\hat{f}^0(X) = \bar{y}$ , hvor  $\bar{y}$  angiver gennemsnittet af  $y$ . Beregn residualerne  $r_j = y_j - \hat{f}^0(X_j)$  for alle observationer  $j = 1, \dots, n$ .
2. For  $b = 1, 2, \dots, B$ :
  - a. Dan et træ  $\hat{f}^b$  med  $d$  split ( $d+1$  terminalnoder) på datasættet  $\{(X_j, r_j)\}_{j=1}^n$ .
  - b. Opdatér modellen med en nedskaleret version af det nye træ:  $\hat{f}(X) \leftarrow \hat{f}(X) + \lambda \cdot \hat{f}^b(X)$
  - c. Opdatér residualerne:  $r_j \leftarrow r_j - \lambda \cdot \hat{f}^b(X_j)$
3. Anvend den boostede model:

$$\hat{f}(X) = \hat{f}^0(X) + \sum_{i=1}^B \lambda \cdot \hat{f}^i(X)$$

Anm.: Beskrivelsen tager udgangspunkt i boostingprocessen anvendt til regression. Metoden til klassifikationsproblemer er tilsvarende men mere kompleks.

Kilde: Egen tilblivelse på baggrund af James, Witten, Hastie & Tibshirani (2021).

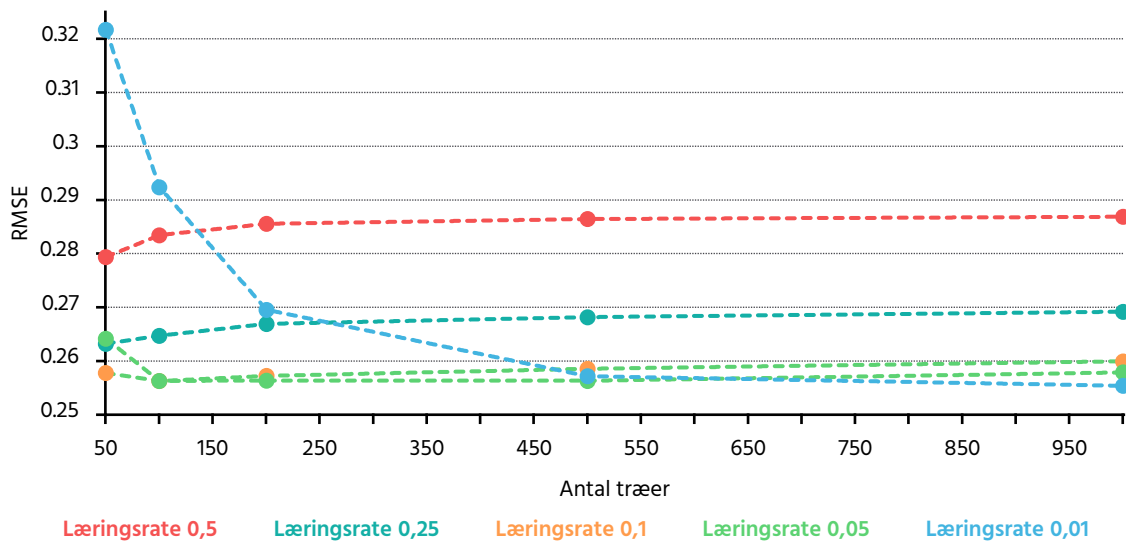
Figur 5.6 viser et eksempel af hyperparameteroptimeringen. Det ses, at hvis få træer inkluderes i modellen, vil relativt høje læringsrater give den bedste prædiktionssevne. Modsat forholdet sig ved et højt antal træer.

I nærværende tilfælde skal afvikles et meget stort antal estimationer. Derfor inddrages estimationens tidsforbrug ligeledes i afvejningen, når det anvendte sæt af hyperparametre bestemmes. Det vil for eksempel ikke være praktisk muligt at afvikle en model med en lav læringsrate og flere tusind træer 1,4 mio. gange.

En læringsrate i intervallet 0,06-0,1 og antal træer mellem 200-400 er undersøgt nærmere. Det findes, at en model med læringsrate på 0,1 og op til 300 træer giver en lav prædiktionsfejl, hvorfor dette vælges som et formålstjenligt sæt af hyperparametre.



Figur 5.6  
Tuning af hyperparametre



Anm.: Figuren viser LightGBM estimeret for 2.500 tilfældigt udvalgte boliger handlet i perioden 1992-2021. Modellen kører med forskellige værdier af antal træer og læringsraten. For hver model beregnes fejlen ved RMSE mellem den prædikterede og den faktiske handelspris.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

## Overfitting

Et problem i forhold til at anvende en kompleks træbaseret model kan være overfitting. At overfitte data betyder, at modellen tilpasser sig den støj, som observeres i netop det datasæt, modellen estimeres på. Dette er en uønsket situation, idet en sådan model ikke vil give nøjagtige estimater af responsvariablen, når modellen anvendes på nye observationer, der ikke var en del af det oprindelige datasæt.

Hvorvidt en model overfitter data tjekkes sædvanligvis ved at opdele data i to dele, hvor modellen estimeres på den ene del (kaldet træningsdata), indeholder måske 80 pct. af observationerne. Modellens prædiktionssevne overfor fremmed data testes på de resterende observationer (kaldet testdata)<sup>16</sup>.

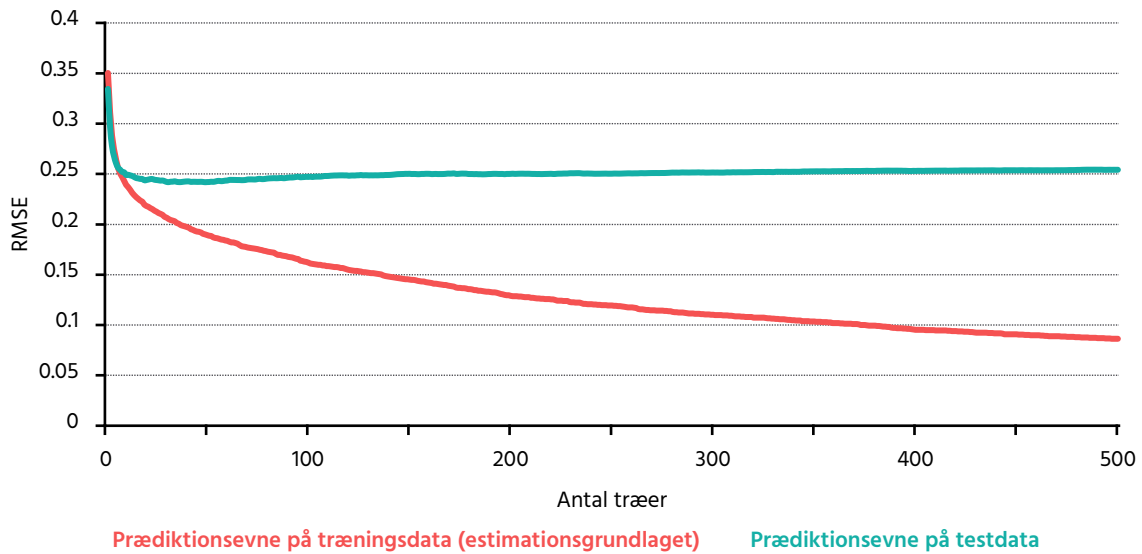
Figur 5.7 viser et eksempel på overfitting. Det ses at modellen estimeret på træningsdata vedvarende opnår lavere fejl, når antallet af træer øges. Når den resulterende model testes på testdata, så begynder fejlen at stige, når modellen indeholder flere end cirka 65 træer.

I nærværende praktiske anvendelse af LightGBM til at bestemme en boligs kvadratmeterpris er der kun i få tilfælde observeret tendens til overfitting, så længe modellens læringsrate fastsættes forholdsvis lavt på 0,1 eller lavere. Overfitting vurderes på denne baggrund som et forholdsvis begrænset problem, som dog har medført høj prædiktionsfejl på få boliger. Det er valgt at tage hensyn til overfitting, ligesom den valgte implementering mindsker afviklingstiden.

<sup>16</sup> Ligeledes kan krydsvalidering (eng: cross validation) anvendes, hvilket oftest vil være bedre end blot én opdeling af data. Fremgangsmåden i krydsvalidering er at dele data op i  $n$  dele. Der foretages efterfølgende  $n$  estimationer af modellen, hvor de forskellige dele af data på skift anvendes til estimation af modellen ( $n-1$  dele af data) og vurdering af modellens nøjagtighed (den resterende del).

Figur 5.7

Eksempel på overfitting og tidlig afbrydelse



Anm.: Figuren viser estimation af 500 modeller, hvor antal træer varieres. Læringsraten antages lig 0,3 i alle modeller. For hver model udregnes prædiktionsvnen på både trænings- og testdata. Modellen er kørt på datagrundlag indeholdende de 10.000 nærmeste bolighandler til et parcelhus i Odense. Træningsdata indeholder 8.000 tilfældigt udvalgte boliger, mens testdata indeholder de resterende 2.000.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Implementeringen af LightGBM indeholder en metode til at undgå overfitting<sup>17</sup>. Ideen er, at estimationens datagrundlag opsplittes i trænings- og testdata. Hvert træ i boosting proceduren estimeres på træningsdata, og prædiktionsvnen i modellen opdateret med dette træ udregnes på testdata. Der anvendes vægtet test, så modellen optimeres til at prædiktere boligprisen tæt ved den betragtede bolig. Hvis ikke modellens prædiktionssevne forbedres set over de seneste  $e$  iterationer, så afbrydes boosting processen, så ikke modellen begynder at overfitte data. Dette kaldes tidlig afbrydelse (eng: early stopping). Parameteren  $e$  sættes af modelbrugeren. En tommelfingerregel siger, at  $e$  bør være i størrelsesordenen 10 pct. af det anvendte antal træer.

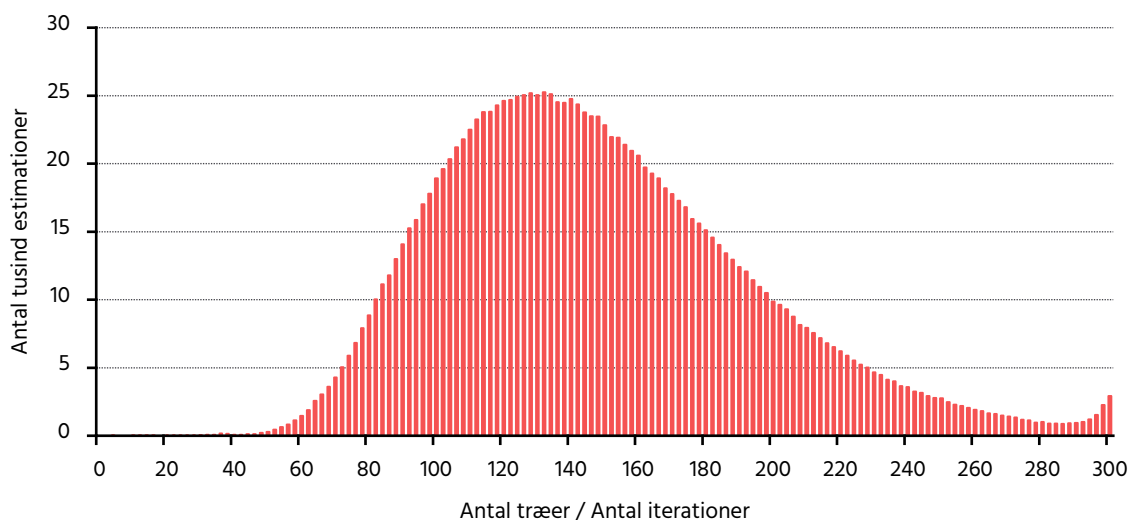
Ved eksemplet illustreret i Figur 5.7 vil tidlig afbrydelse af algoritmen betyde, at metoden stopper efter er cirka 115 iterationer. Metoden afbrydes her, da den optimale model med mindst fejl evalueret på testdata findes efter cirka 65 iterationer og stopkriteriet typisk vil være, at der ikke opnås forbedring i cirka 50 iterationer (10 pct. af i alt 500 træer). Modellen med mindst prædiktionsfejl på testdata vil blive anvendt, også selvom der er foretaget yderligere iterationer sidenhen.

<sup>17</sup> Ud over tidlig afbrydelse indeholder træbaserede algoritmer typisk en række parametre, der fastsætter træstrukturen i modellen. Værdien af disse parametre kan ligeledes forsøges justeret, hvis man oplever tendens til, at modellen overfitter. Træerne kan gøres simple, og dermed have lavere tilbøjelighed til at overfitte, hvis man for eksempel (1) sætter et minimum antal observationer i hver terminalnode, (2) en maksimal dybde for hvert træ, (3) estimerer hvert enkelt træ på et unikt datasæt i form af en tilfældigt udvalgt stikprøve af træningsdata (for eksempel kun en andel af rækker eller søjler, som skifter fra træ til træ), (4) ændre antal kategorier for en kontinuert variabel (LightGBM inddeler en kontinuert variabel i en række grupper, antal grupper kan sænkes for at få en simple model). Modeller for beslutningstræer indeholder typisk mange af sådanne parametre, som kan tunes.

Det er valgt at anvende tidlig afbrydelse, så overfitting minimeres. Idet der anvendes en model med et relativt stort antal træer, som skal afvikles mange gange, vil tidlig afbrydelse desuden spare unødigt beregningstid, idet læringsprocessen afbrydes, når modellens prædiktionssevne ikke længere forbedres.

I praksis indgår der i gennemsnit knap 150 træer i den estimerede model. Tidlig afbrydelse nedbringer således afviklingstiden til cirka det halve i forhold til, hvis samtlige estimationer skulle afvikle de maksimale 300 træer. Figur 5.8 viser fordelingen af faktisk antal træer, som indgår i hver estimation. Størstedelen af estimationerne afbrydes inden, det maksimale antal træer nås.

**Figur 5.8**  
Antal iterationer



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

## 5.5 Estimationsresultater og træfsikkerhed

For samtlige 1,4 mio. ejerboliger i Danmark imputeres boligprisen for alle kvartaler i perioden 1992-2021 med udgangspunkt i geografisk vægtet regression, hvor regressionsanalysen udføres med den træbaserede maskinlæringsalgoritme LightGBM. De indeholdte baggrundskarakteristika på boligen og dens ejere er angivet i Tabel 5.4 på side 82. For hvert geografisk område baseres analysen på det antal boliger, som de optimerede søgeområdeparametre giver anledning til.

Det store antal estimationer giver et stort antal estimerede parametre, ligesom parameter-værdierne kan variere på tværs af landet. Det er derfor ikke muligt at afbilde samtlige parameterestimer.

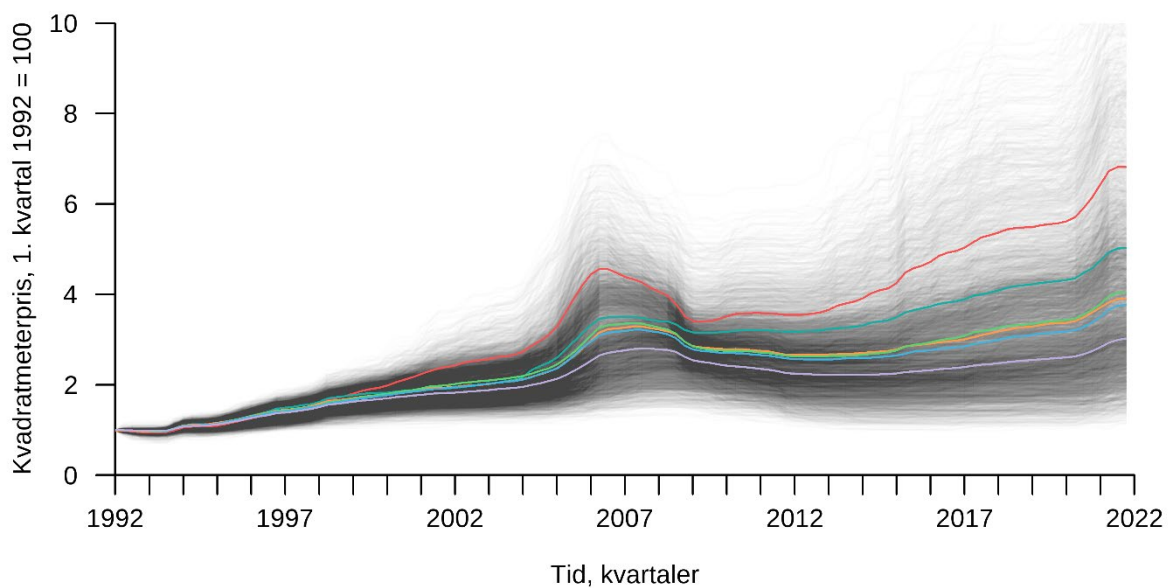
### Estimationsresultater

Et skøn over prisudviklingen for ejerboliger i perioden siden 1992 er metodens hovedresultat. Figur 5.9 viser det skønnede prisindeks for 5.000 tilfældigt udvalgte ejerboliger i hvert kvartal i perioden 1992-2021. Derudover er der indtegnet det gennemsnitlige prisindeks for forskellige bystørrelser. Prisindeksenes referencekvartal er sat til første kvartal i den betragtede periode.

Fra figuren ses midt 2000'ernes boligprisboble med kraftige prisstigninger fra starten af perioden til omkring 2007. Herefter vendte udviklingen og ejerboligprisen faldt. Siden start 2010'erne har boligmarkedet igen været i fremgang.

Særligt i hovedstadsområdet og i de større byer er boligpriserne volatile, idet disse områder oplever de største prisstigninger og -fald. Der ses en tendens til, at bevægelser i ejerboligpriserne i hovedstadsområdet over tid forplantes ned igennem de forskellige bystørrelser, hvor større byer påvirkes først og mest, mens de mindste byer påvirkes sidst og mindst. Denne effekt er ofte benævnt bølgeeffekten (eng: the ripple effect), som netop beskriver den pris-drevne regionale substitutionseffekt, jf. Hviid et al. (2016) samt Meen (1999) og (2001).

**Figur 5.9**  
Prisindeks for 5.000 tilfældigt udvalgte ejerboliger



Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

### Træfsikkerhed

Modellens træfsikkerhed måles dels på 10 pct. af samtlige bolighandler i den betragtede periode. Her prædikteres med faktiske karakteristika på både boligen og dens ejere, hvorefter afvigelsen mellem den af modellen prædikterede kvadratmeterpris og den faktisk realiserede udregnes.

Derudover kan træfsikkerheden i den enkelte estimation vurderes ved at betragte modellens prædiktionssevne på de 3.000 boliger i testdata. Dette er dog en mere upræcis test, som dog alligevel angives i Figur 5.10. I denne figur indgår alle boliger, ikke kun dem som er handlet inden for den betragtede periode.

Den gennemsnitlige afvigelse i de prædikterede kvadratmeterpriser er 1.732 kr. svarende til 16,5 pct., jf. Tabel 5.2. For tabellens fem øverste rækker gælder, at træfsikkerheden er størst, når afvigelsen er så lille som mulig (minimum nul). For de tre nederste rækker skal værdien være størst mulig (maksimalt 100). Som beskrevet i afsnit 5.3 giver geografiske modeller baseret på maskinel læring den højeste træfsikkerhed. I nederste halvdel af tabellen er desuden angivet træfsikkerhed for den tilsvarende model med lineær regression.

Kvadratrod af den gennemsnitlig kvadratafgivelse (eng: root-mean-square error, RMSE) angiver standardafvigelsen mellem den prædikterede og faktiske kvadratmeterpris (residualerne). Dette mål er et estimat for spredningen på residualerne og er derfor ofte anvendt. Som supplement hertil angives, hvor stor en andel af det samlede antal bolighandler, hvor den prædikterede kvadratmeterpris ligger indenfor +/- 20 pct. af den faktiske handelspris (kaldet PM20). At udføre regressionen med maskinel læring øger PM20 med 2,7 pct. point i forhold til den tilsvarende model udført med traditionel regression.

**Tabel 5.2**  
 Træfsikkerhed i geografisk vægtet regression ved diverse fejlmål

	Alle boliger	Huse	Lejligheder
<b>Træfsikkerhed ved maskinlæring (LightGBM)</b>			
Gennemsnitlig afvigelse (MAE)	1.732	1.730	1.739
Gennemsnitlig kvadratafgivelse (MSE)	9.340.514	9.509.278	8.816.163
Gennemsnitlig procentvis afvigelse (MAPE)	16,5	18,2	11,3
Kvadratrod af gennemsnitlig kvadratafgivelse, m <sup>2</sup> -pris (RMSE)	3.056,2	3.083,7	2.969,2
Kvadratrod af gennemsnitlig kvadratafgivelse, log-m <sup>2</sup> -pris (RMSE)	0,232	0,250	0,159
Andel i pct. som rammer indenfor +/- 20 pct. (PM20)	76,1	72,4	87,6
Andel i pct. som rammer indenfor +/- 10 pct. (PM10)	49,5	45,4	62,3
Andel i pct. som rammer indenfor +/- 5 pct. (PM5)	27,2	24,5	35,6
<b>Træfsikkerhed ved regression (OLS)</b>			
Gennemsnitlig afvigelse (MAE)	1.876	1.837	1.999
Gennemsnitlig kvadratafgivelse (MSE)	10.828.030	10.827.339	10.830.179
Gennemsnitlig procentvis afvigelse (MAPE)	18,0	19,6	13,0
Kvadratrod af gennemsnitlig kvadratafgivelse, m <sup>2</sup> -pris (RMSE)	3.290,6	3.290,5	3.290,9
Kvadratrod af gennemsnitlig kvadratafgivelse, log-m <sup>2</sup> -pris (RMSE)	0,246	0,264	0,178
Andel i pct. som rammer indenfor +/- 20 pct. (PM20)	73,4	70,2	83,2
Andel i pct. som rammer indenfor +/- 10 pct. (PM10)	45,5	42,5	54,6
Andel i pct. som rammer indenfor +/- 5 pct. (PM5)	24,3	22,5	29,9

Anm.: Træfsikkerheden er målt for 10 pct. af de handlede boliger i perioden 1992-2021. Her sammenholdes den faktiske handelspris med modellens prædikterede pris, idet prædiktionen foretages med boligens og boligejerens faktiske karakteristika på handelstidspunktet.

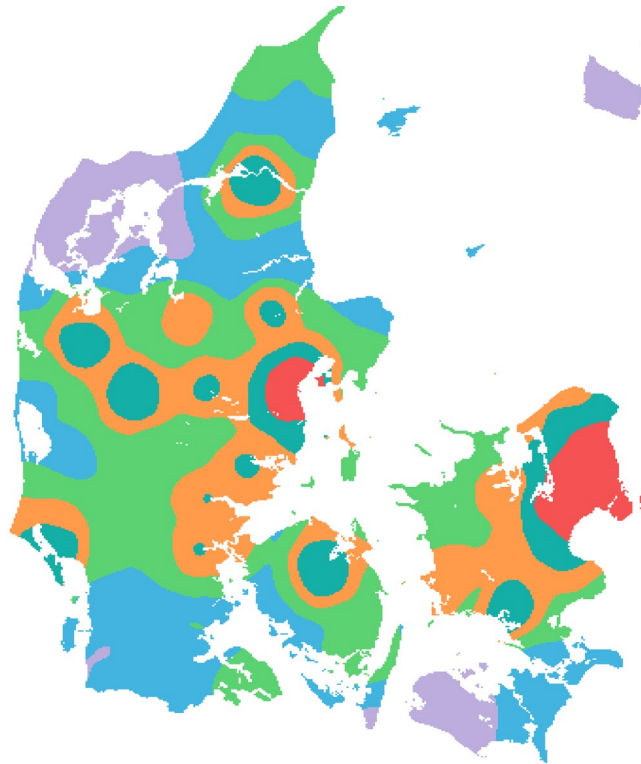
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Generelt opnås den største træfsikkerhed i de større byområder. Det fremgår af Figur 5.10, der viser den geografiske variation i PM20. Det virker naturligt, at modellen bedst vurderer ejendomme i de større byer, hvor der dels sælges flest boliger og boligmassen dels er mere homogen, end det er tilfældet i landdistrikterne. Områderne med den laveste træfsikkerhed ses at være de mindre øsamfund, Lolland, Falster, Sydvestjylland og Nordvestjylland.

Figur 5.10

Træfsikkerhed i geografisk vægtet regression

---



PM20 op til 59 pct. PM20 59-65 pct. PM20 65-71½ pct. PM20 71½-77 pct. PM20 77-84 pct. PM20 84-98 pct.

Anm.: Træfsikkerheden er målt for hver estimation ved at teste modellens prædiktionssevne på de 3.000 boliger i testdata. Det måles, hvor stor en andel af de estimerede ejendomsværdier, som ligger indenfor +/- 20 pct. af den faktiske handelspris (PM20).

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

At anvende maskinlæring frem for lineær regression har øget modellens prædiktionssevne i alle landsdele, jf. Tabel 5.3. Målt ved PM20 er den største forbedring sket omkring København og på Bornholm.

**Tabel 5.3**

Træsikkerhed i geografisk vægtet regression efter landsdel, bystørrelse og boligtype

	Kvadratrod af gns. kvadratafvigelse (RMSE)		Gennemsnitlig procentvis afvigelse (MAPE)		Andel med pris inden for +/- 20 pct. (PM20)	
	OLS	LightGBM	OLS	LightGBM	OLS	LightGBM
<b>Hele landet</b>						
Alle ejerboliger	0,246	0,232	18,0	16,5	73,4	76,1
Huse	0,264	0,250	19,6	18,2	70,2	72,4
Ejerlejligheder	0,178	0,159	13,0	11,3	83,2	87,6
<b>Efter landsdel, huse</b>						
Byen København	0,202	0,190	15,4	14,2	77,1	80,3
Københavns omegn	0,185	0,175	13,5	12,6	82,2	84,3
Nordsjælland	0,204	0,204	14,9	14,5	78,9	78,8
Bornholm	0,331	0,309	26,7	24,2	55,3	57,5
Østsjælland	0,201	0,193	14,7	13,7	80,0	81,2
Vest- og Sydsjælland	0,280	0,273	21,5	20,5	66,2	68,2
Fyn	0,275	0,260	20,7	19,0	67,7	70,6
Syddjylland	0,287	0,270	21,8	20,0	66,8	69,7
Østjylland	0,254	0,244	18,6	17,6	72,7	73,8
Vestjylland	0,288	0,266	21,5	19,6	66,1	69,3
Nordjylland	0,309	0,287	23,8	21,5	61,9	65,6
<b>Efter landsdel, ejerlejligheder</b>						
Byen København	0,170	0,164	12,3	11,6	85,7	87,3
Københavns omegn	0,171	0,147	12,7	10,3	84,1	90,7
Nordsjælland	0,180	0,160	13,4	11,2	80,6	85,9
Bornholm	0,257	0,209	24,5	17,5	44,0	60,0
Østsjælland	0,178	0,158	12,5	10,3	86,9	92,2
Vest- og Sydsjælland	0,199	0,156	14,7	11,5	78,5	86,4
Fyn	0,197	0,180	14,6	12,8	80,4	85,4
Syddjylland	0,213	0,174	16,0	12,7	74,8	82,9
Østjylland	0,168	0,148	12,3	10,6	84,3	88,7
Vestjylland	0,206	0,163	15,6	11,9	75,0	85,7
Nordjylland	0,188	0,162	14,2	11,7	79,0	85,9
<b>Efter bystørrelse ved antal indbyggere, huse</b>						
Hovedstadsområdet	0,189	0,178	14,0	13,0	80,8	83,1
Mindst 100.000	0,203	0,189	15,0	13,5	79,0	82,0
20.000-99.999	0,218	0,206	16,0	14,8	76,2	78,6
10.000-19.999	0,227	0,215	16,8	15,7	74,1	76,8
5.000-9.999	0,234	0,221	17,2	16,1	73,4	76,5
1.000-4.999	0,251	0,239	19,1	17,8	70,6	72,9
Op til 999	0,364	0,346	29,3	27,3	54,3	55,3

	Kvadratrod af gns. kvadratafvigelse (RMSE)		Gennemsnitlig procentvis afvigelse (MAPE)		Andel med pris inden for +/- 20 pct. (PM20)	
	OLS	LightGBM	OLS	LightGBM	OLS	LightGBM
<b>Efter bystørrelse ved antal indbyggere, ejerlejligheder</b>						
Hovedstadsområdet	0,172	0,160	12,5	11,4	85,1	88,0
Mindst 100.000	0,163	0,146	12,1	10,3	85,5	89,8
20.000-99.999	0,186	0,158	13,8	11,2	80,8	87,0
10.000-19.999	0,209	0,171	15,3	12,3	75,3	84,4
5.000-9.999	0,190	0,166	14,4	12,3	75,9	84,5
1.000-4.999	0,245	0,209	19,2	15,8	71,6	78,4
Op til 999	0,222	0,215	17,8	15,8	67,8	76,1

Anm.: Træfsikkerheden er målt for 10 pct. af de handlede boliger i perioden 1992-2021. Her sammenholdes den faktiske handelspris med modellens prædikterede pris, idet prædiktionen foretages med boligens og boligejerens faktiske karakteristika på handelstidspunktet.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

## Variablernes betydning

En ulempe ved træbaserede modeller baseret på mange beslutningstræer kan være, at det er vanskeligt at tolke og visualisere den anvendte model.

For eksempel er det umiddelbart svært at afgøre, hvilke af de forklarende variable, som har størst betydning for en ejerboligs kvadratmeterpris. Algoritmen kan dog rangere de indgående variable efter deres betydning for det analyserede udfald. Ideen ved rangeringen er, at hvis en variabel ikke er vigtig, så vil det kun have lille effekt på modellens prædikterede kvadratmeterpriser, hvis variabelens værdi ændres.

Figur 5.11 viser vigtigheden af de enkelte variable i den estimerede model. Jo større værdi, jo vigtigere er den enkelte variabel for fastsættelsen af en ejerboligs kvadratmeterpris. Vurderet ud fra dette indeks er den vigtigste forudsætning for boligprisen, hvornår i perioden boligen er handlet. Dette virker naturligt, da boligpriserne over perioden oplever store tidsmæssige udsving, for eksempel prisstigningerne i perioden frem mod 2007 og det efterfølgende fald.

Herefter følger karakteristika på boligen i form af opførelsesår, grundstørrelse (for ikke-lejligheder) og boligens samlede areal. En boligs opførelsesår vil ofte udtrykke en grad af attraktivitet og byggekvalitet som for eksempel muremestervilla, typehus fra 1960'erne osv. Opførelsesår kan ligeledes indikere boligens isoleringsforhold, som kan påvirke handelsprisen. Inden for et større byområde kan derudover være en tendens til, at ældre boliger er attraktivt beliggende tæt ved centrum, mens parcelhuse opført senere hen kan ligge i større afstand til bycentrum.

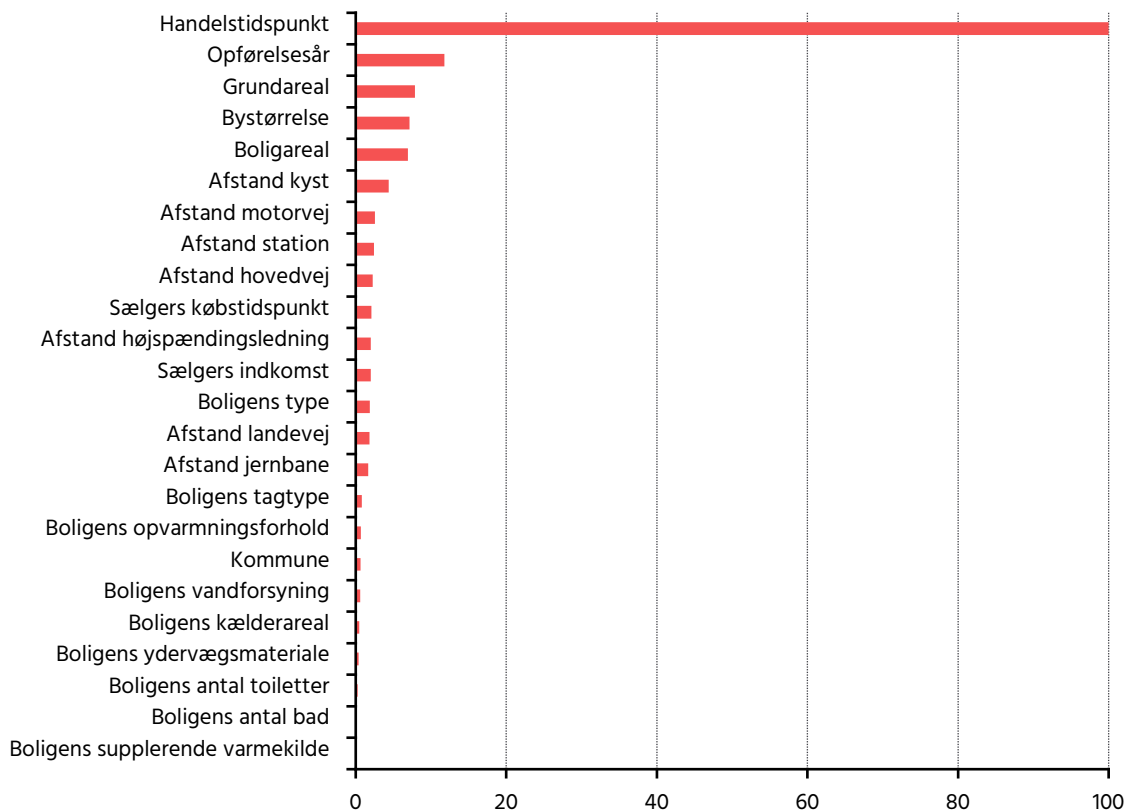
Boligens beliggenhed har også betydning for prisen, hvor særligt bystørrelse, afstand til kyst, station og veje påvirker kvadratmeterprisen.

Rangeringen angivet i figuren nedenfor er udregnet som gennemsnit for alle boliger i hele landet og udtrykker kun, at et karakteristika ofte påvirker ejerboligprisen. Indekset afslører ikke, hvor meget et ændret karakteristika ændrer skønnet for ejendomsværdien.



Figur 5.11

Forklarende variabelers betydning for boligens kvadratmeterpris



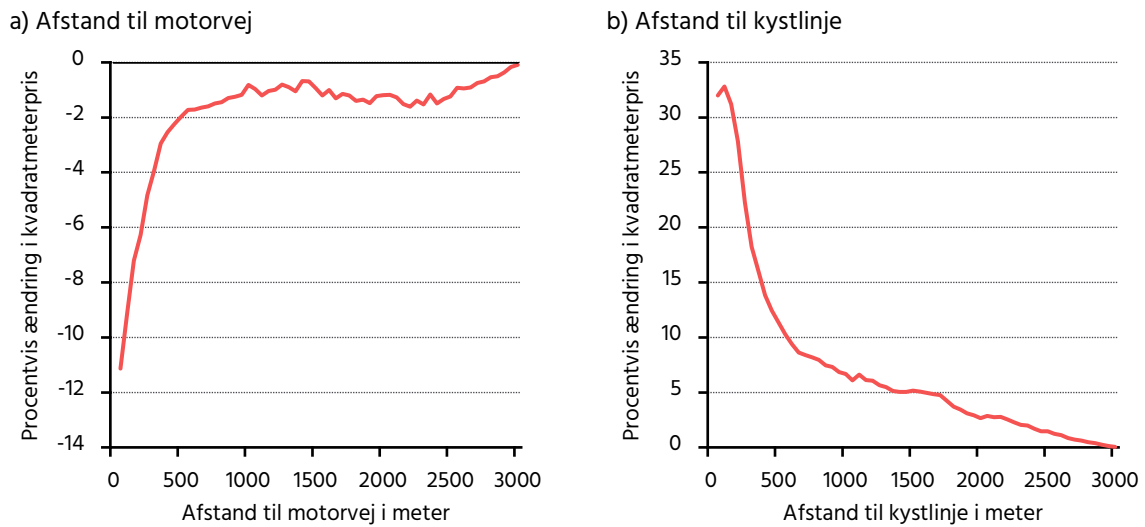
Anm.: Variablenes 'vigtighed' er normeret i forhold til den bedst forklarende variabel. Udfaldsrum for de enkelte variable kan ses i Tabel 5.4.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Som eksempel på betydningen af de anvendte karakteristika for boligens beliggenhed betragtes de to variable, som angiver afstand til henholdsvis motorvej og kyst, jf. Figur 5.12. Det betragtes, hvordan nærhed til motorvej og kystlinje påvirker kvadratmeterprisen i forhold til den hypotetiske situation, at selvsamme bolig i stedet var beliggende 3.000 meter fra henholdsvis motorvejen og kystlinjen.

Modellen angiver, at beliggenhed op ad motorvej skønnes at reducere kvadratmeterprisen med op mod 12 pct. Modsat øges kvadratmeterprisen med op til en tredjedel, hvis ejerboligen er beliggende ud til kysten. For begge karakteristika gælder, at priseffekten aftager kraftigt til en afstand på omkring 600 meter. Herefter aftager effekten i svagere takt frem mod en afstand på 3.000 meter, hvor priseffekten pr. definition er nul.

**Figur 5.12**  
 Effekt af afstand til motorvej og kyst



Anm.: Figuren viser den estimerede effekt på prisen pr. kvadratmeter ved boligens faktiske placering i forhold til den skønnede kvadratmeterpris, hvis præcis samme bolig var beliggende 3.000 meter fra henholdsvis motorvej og kystlinje.  
 Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

## 5.6 Forklarende variable

Tabel 5.4 indeholder en oversigt over de baggrundsvariable, som inkluderes i de estimerede modeller. Tabellens midterste kolonne viser variablene anvendt ved træbaserede algoritmer, hvor kontinuerte variable kan inkluderes. Ved simpel regression vil en kontinuert variabel ofte blive inddelt i intervaller, så modellen tillades at fastsætte en ikke-lineær sammenhæng mellem boligprisen og for eksempel afstand til nærmeste kystlinje. Kolonnen yderst til højre angiver de anvendte intervalstørrelser.

**Tabel 5.4**  
 Forklarende variable anvendt i geografisk vægtet regression

Variabel	Udfaldsrum ved maskinlæring (LightGBM)	Udfaldsrum ved regression (OLS)
Konstantled	-	Konstant
Kommunenavn	Kategorisk: 98 kommuner	
<b>Geografiske karakteristika: Afstand til nærmeste</b>		
Kystlinje	Kontinuert: Antal meter	Kategorisk: {0-200 m.; 200-1.500 m.; 1.500-5.000 m.; 5.000-9.000 m.; over 9.000 m.}
Højspændingsledning	Kontinuert: Antal meter	Kategorisk: {0-1 km.; 1-2½ km.; 2½-7½ km.; over 7½ km.}
Jernbane	Kontinuert: Antal meter	Kategorisk: {0-200 m.; 200-1.700 m.; 1.700-3.200 m.; 3.200-6.300 m.; over 6.300 m.}
Station	Kontinuert: Antal meter	Kategorisk: {0-1 km.; 1-3 km.; 3-5 km.; 5-10 km.; over 10 km.}
Motorvej	Kontinuert: Antal meter	Kategorisk: {0-0,2 km.; 0,2-2,8 km.; 2,8-4,3 km.; 4,3-7,4 km.; 7,4-20,0 km.; over 20 km.}

Variabel	Udfaldsrum ved maskinlæring (LightGBM)	Udfaldsrum ved regression (OLS)
Hovedvej	Kontinuert: Antal meter	Kategorisk: {0-100 m.; 100-300 m.; 300-1.400 m.; 1.400-3.000 m.; over 3.000 m.}
Landevej	Kontinuert: Antal meter	Kategorisk: {0-100 m.; 100-200 m.; 200-1.900 m.; over 1.900 m.}

#### Boligkarakteristika

Areal	Kontinuert: Antal m <sup>2</sup>	Kategorisk: {0-40 m <sup>2</sup> ; 40-60 m <sup>2</sup> ; 60-80 m <sup>2</sup> ; 80-100 m <sup>2</sup> ; 100-120 m <sup>2</sup> ; 120-140 m <sup>2</sup> ; 140-160 m <sup>2</sup> ; 160-180 m <sup>2</sup> ; 180-200 m <sup>2</sup> ; 200-220 m <sup>2</sup> ; 220-240 m <sup>2</sup> ; 240-260 m <sup>2</sup> ; 260-280 m <sup>2</sup> ; 280-300 m <sup>2</sup> ; over end 300 m <sup>2</sup> }
Kælderareal	Kontinuert: Antal m <sup>2</sup>	Kategorisk: {0m <sup>2</sup> ; 1-50m <sup>2</sup> ; 51-150m <sup>2</sup> ; mere end 150 m <sup>2</sup> }
Grundareal	Kontinuert: Antal m <sup>2</sup>	Kategorisk: {0-200m <sup>2</sup> ; 201-700m <sup>2</sup> ; 701-1200m <sup>2</sup> ; over 1200 m <sup>2</sup> }
Opførelsesår	Kontinuert: Årstal	Kategorisk: {Opført før 1800; opført 1800-1889; 1890-1899; 1900-1909; 1910-1919; 1920-1929; 1930-1939; 1940-1949; 1950-1959; 1960-1969; 1970-1979; 1980-1989; 1990-1999; 2000-2009; efter 2010}
Boligtype	Kategorisk: stuehus, parcelhus, rækkehus eller lejlighed.	
Ydervægsmateriale	Kategorisk: mursten mv., beton mv., bindingsværk, træbeklædning	
Tagdækningsmateriale	Kategorisk: fladt tag, tagpap, fibercement, cementsten, tegl mv., metalplader, stråtag	
Primær opvarmningsform	Kategorisk: el, olie, fast brændsel, naturgas, fjernvarme, varmepumpe, andet	
Supplerende varme	Dummy: lig 1 hvis boligen har brændeovn eller pejs, 0 ellers	
Vandforsyning	Kategorisk: Offentligt alment, privat alment, egen boring, brønd, ingen vandforsyning	
Antal toiletter	Kontinuert: Antal toiletter	Kategorisk: 1,2, 3 eller flere
Antal bad	Kontinuert: Antal badeværelser	Dummy: Lig én mere end et bad, 0 ellers

#### Ejerkarakteristika (karakteristika på sælger, når hvis boligen sælges i løbet af året)

Indkomst	Kategorisk: Indkomstdecil for husstandsækvivaleret disponibel indkomst	
År siden køb	Kontinuert: År	-
Indkomst $\times$ År siden køb	-	Interaktionsled mellem indkomstdecil og antal år siden køb som kategorisk variabel

#### Tid

Handelskvartal		Kategorisk: Salgstidspunkt opgjort som kvartal for salg, for eksempel 2. kvartal 2013
Handelsdag	Kontinuert: Decimaltal, f.eks. 5. maj 2013 = 2013,3425	

Anm.: For ejerlejligheder antages både grundareal og kælderareal at være nul.



## 6. Referencer

Arya, S., Mount, D. M., Netanyahu, N. S., & Silverman, R. o. (1998). An Optimal Algorithm for Approximate Nearest Neighbor Searching in Fixed Dimensions. *Journal of the ACM*, vol. 45, nr. 6, s. 891–923.

Arya, S., Mount, D., Kemp, S. E., & Jefferis, G. (1. august 2023). *Package 'RANN': Fast Nearest Neighbour Search (Wraps ANN Library) Using L2 Metric*. Hentet fra The Comprehensive R Archive Networks hjemmeside: <https://cran.r-project.org/web/packages/RANN/RANN.pdf>

Boligøkonomisk Videncenter. (2019). *Danske boligprisindeks 1938-2017 samt historiske data om boligmarkedet*. København: Boligøkonomisk Videncenter.

Boligøkonomisk Videncenter. (1. august 2023). *Boligprisindeks*. Hentet fra Boligøkonomisk Videncenters hjemmeside: <https://www.bvc.dk/tal-statistik/boligprisindeks/>

Brunsdon, C., Fotheringham, A., & Charlton, M. (oktober 1966). Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis* 28(4), s. 281–298.

Chen, L., Ren, C., Zhang, B., Wang, Z., & Xi, Y. (2018). Estimation of Forest Above-Ground Biomass by Geographically Weighted Regression and Machine Learning with Sentinel Imagery. *Forests* 9(10), 582.

Chen, T., & Guestrin, C. (august 2016). Xgboost: A scalable tree boosting system. *22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining*, s. 785–794.

Court, A. T. (1939). Hedonic Price Indexes. *The Dynamics of Automobile Demand, General Motors Corporation*, s. 99-119.

Danmarks Statistik. (8. april 2015). *Flere bor til leje - Boligopgørelsen 1. januar 2015*. Hentet fra Danmarks Statistiks hjemmeside: <https://www.dst.dk/da/Statistik/nyheder-analyser-publ/nyt/NytHtml?cid=19472>

Danmarks Statistik. (2018). *Inddeling af Danmarks kommuner*. København: Danmarks Statistik.

Danmarks Statistik. (1. august 2023). *Statistikdokumentation: Boligopgørelsen*. Hentet fra Danmarks Statistiks hjemmeside: <https://www.dst.dk/da/Statistik/dokumentation/statistikdokumentation/boligopgoerelsen>

Danmarks Statistik. (1. august 2023). *Statistikdokumentation: Byopgørelsen*. Hentet fra Danmarks Statistiks hjemmeside: <https://www.dst.dk/da/Statistik/dokumentation/statistikdokumentation/byopgoerelsen>

Danmarks Statistik. (1. august 2023). *Statistikdokumentation: Ejendomssalg*. Hentet fra Danmarks Statistiks hjemmeside: <https://www.dst.dk/da/Statistik/dokumentation/statistikdokumentation/ejendomssalg>

De Økonomiske Råd. (2016). *Dansk Økonomi, efterår 2026 - Kapitel V: Indkomst- og formuefordeling*. København: De Økonomiske Råd.

Finans Danmark. (1. august 2023). *Datagrundlaget for Boligmarkedsstatistikken*. Hentet fra Finans Danmarks hjemmeside: <https://finansdanmark.dk/tal-og-data/boligstatistik/boligmarkedsstatistikken/datagrundlaget-for-statistikken/>

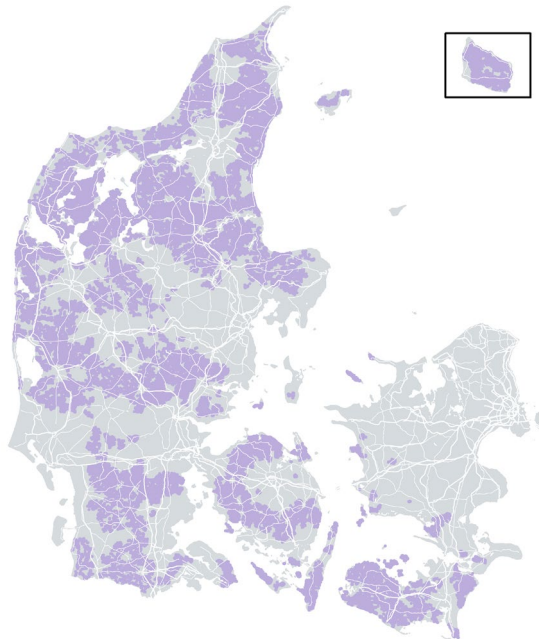
- Hansen, J. Z., Iversen, A. Ø., & Stephensen, P. (2018). *Ejerboliger i det 21. århundrede - En husstandsbaseeret undersøgelse af boligkapitalgevinst og ejerboligbeskatning i perioden 2000–15*. København: DREAM.
- Hviid, S. J., Hvolbøl, T. S., & Pedersen, E. H. (2016). Regionale Aspekter på boligmarkedet. *Danmarks Nationalbank Kvartalsoversigt, 4.kvartal 2016*, s. 47-59.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning : with Applications in R (2nd edition)*. New York: Springer.
- Kampstra, P. (2008). Beanplot: A Boxplot Alternative for Visual Comparison of Distributions. *Journal of Statistical Software, Code Snippets, 28(1)*, s. 1–9.
- Ke, G., Meng, Q., Finley, T., Wang, T., & Chen, W. (december 2017). Lightgbm: A highly efficient gradient boosting decision tree. *31st Conference on Neural Information Processing Systems (NIPS 2017)*, s. 3146–3154.
- Kullback, S., & Leibler, R. A. (1951). On Information and Sufficiency. *The Annals of Mathematical Statistics, 22*, s. 79-86.
- Li, L. (2019). Geographically Weighted Machine Learning and Downscaling for High-Resolution Spatiotemporal Estimations of Wind Speed. *Remote Sensing 11(11)*, 1378.
- Meen, G. (1999). Regional house prices and ripple effect; a new interpretation. *Housing Studies, vol. 14, nr. 6*, s. 733-753.
- Meen, G. (2001). *Modelling spatial housing markets: theory, analysis, and policy - Theory, Analysis and Policy*. Amsterdam: Kluwer Academic Publishers.
- Møller, M. H. (2017). *Boliger som ikke bliver brugt - Parcel- og stuehuse*. København: Boligøkonomisk Videncenter.
- Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogus, A., & Gulin, A. (december 2018). CatBoost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems 31 (NIPS 2018)*, s. 6638–6648.
- Quiñones, S., Goyal, A., & Ahmed, Z. (2021). Geographically Weighted Machine Learning Model for Untangling Spatial Heterogeneity of Type 2 Diabetes Mellitus (T2d) Prevalence in the USA. *Scientific Reports 11(1)*, s. 1–13.
- Rothenberg, J., Galster, G. C., Butler, R. V., & Pitkin, J. R. (1991). *The maze of urban housing markets: Theory, evidence, and policy*. Chicago: University of Chicago Press.
- Stephensen, P., Hansen, J. Z., Jørgensen, E. H., & Hansen, M. F. (2015). *Fremskrivning af ejendomsværdien af parcelhuse*. København: DREAM.
- Xu, S., Zhao, Q., Yin, K., He, G., Zhang, Z., Wang, . . . Zhang, N. (2021). Spatial Downscaling of Land Surface Temperature Based on a Multi-Factor Geographically Weighted Machine Learning Model. *Remote Sensing 13(6)*, 1186.

# A. Ejerboliger efter prisdecil

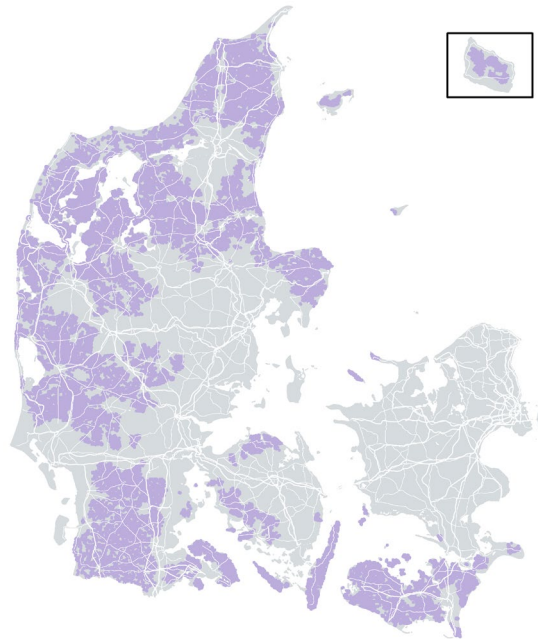
Figur A.1

Ejerboliger med kvadratmeterpris i 1. decil

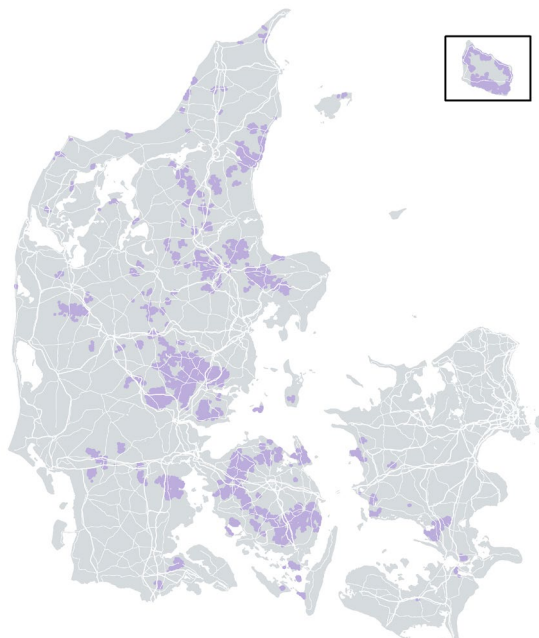
a) 1992



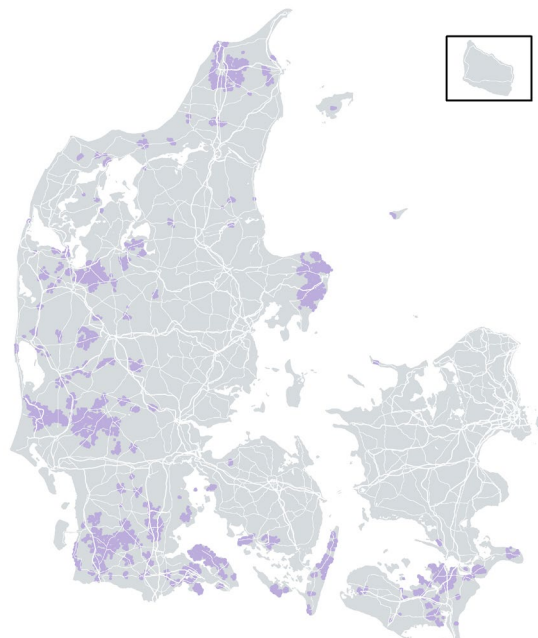
b) 2021



c) Forsvundne områder i 2021



d) Nye områder i 2021



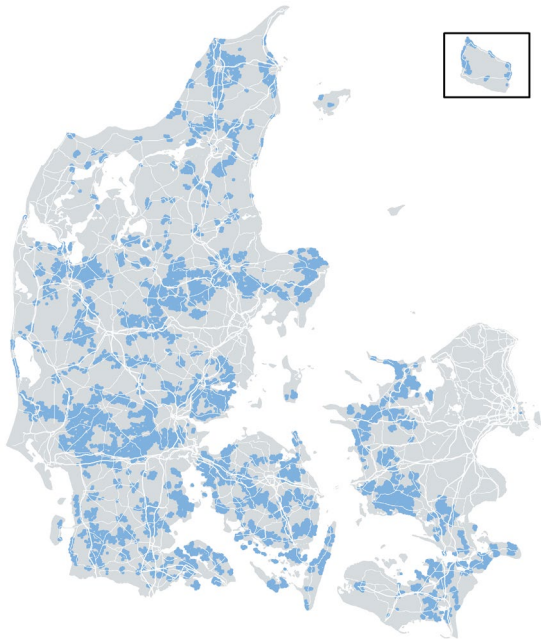
Anm.: Alle ejerboliger er rangeret efter boligens skønnede kvadratmeterpris. Herefter inddeles i ti lige store grupper (deciler) fra 1. decilgruppe indeholder den tiendedel af boligerne med den laveste kvadratmeterpris til 10. decil med de ti pct. af boligerne med den største kvadratmeterpris. Figuren viser placeringen af boliger i 1. decil i både 1992 og 2021. Desuden vises boliger som forlader og tilkommer gruppen i perioden 1992-2021.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

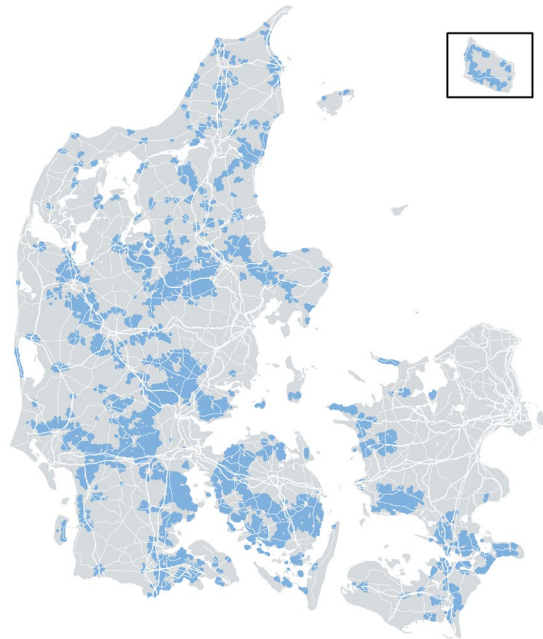
Figur A.2

Ejerboliger med kvadratmeterpris i 2. decil

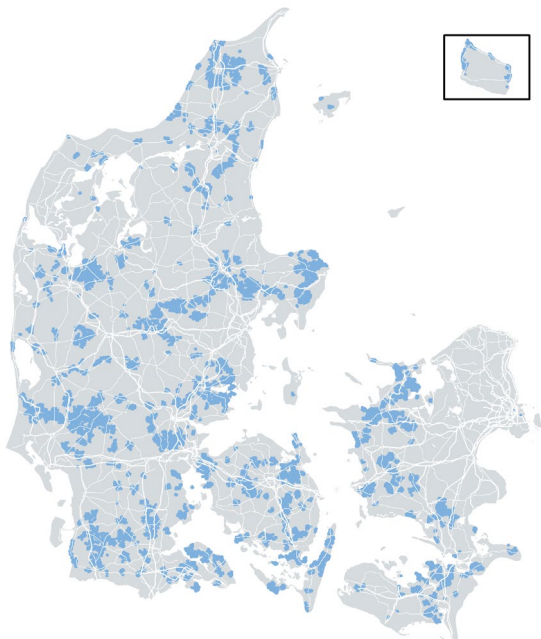
a) 1992



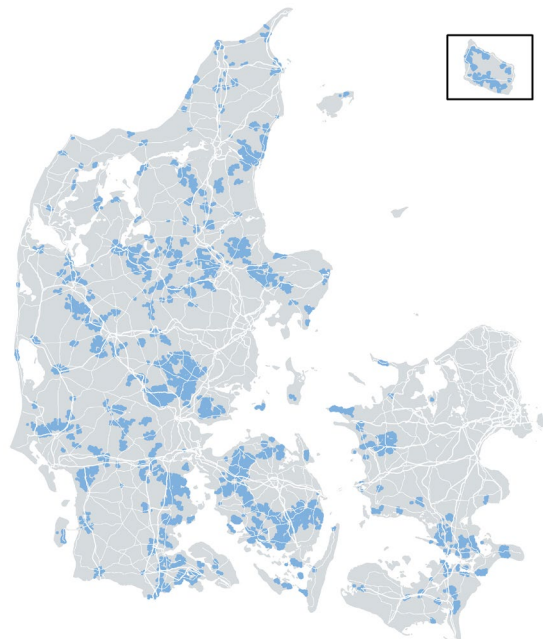
b) 2021



c) Forsvundne områder i 2021



d) Nye områder i 2021



Anm.: Alle ejerboliger er rangeret efter boligens skønnede kvadratmeterpris. Herefter inddeles i ti lige store grupper (deciler) fra 1. decilgruppe indeholder den tiendedel af boligerne med den laveste kvadratmeterpris til 10. decil med de ti pct. af boligerne med den største kvadratmeterpris. Figuren viser placeringen af boliger i 2. decil i både 1992 og 2021. Desuden vises boliger som forlader og tilkommer gruppen i perioden 1992-2021.

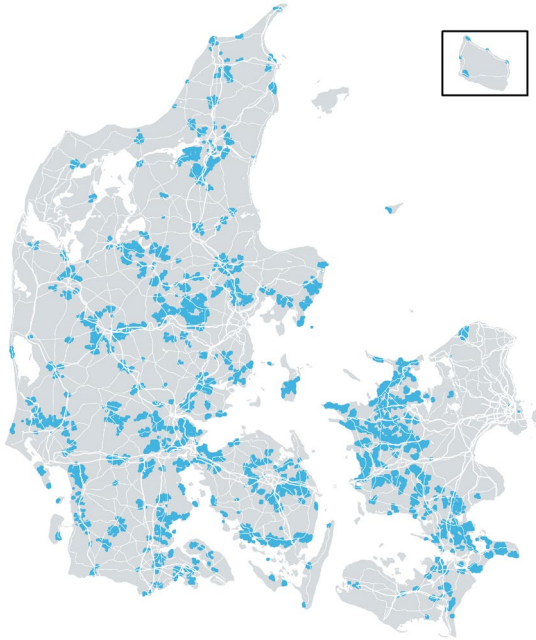
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.



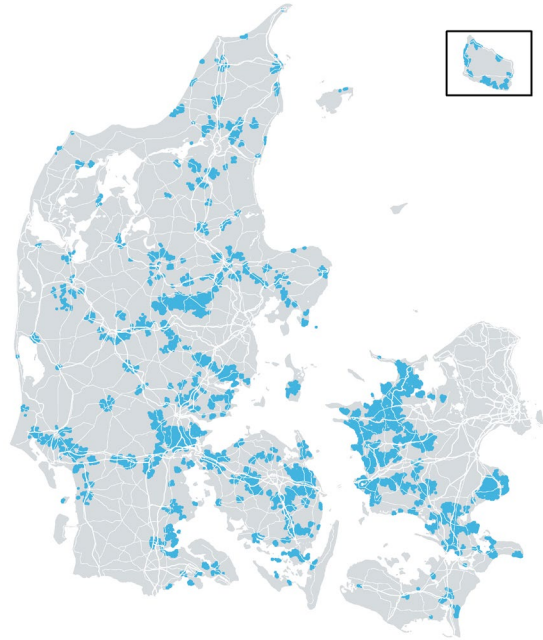
Figur A.3

Ejerboliger med kvadratmeterpris i 3. decil

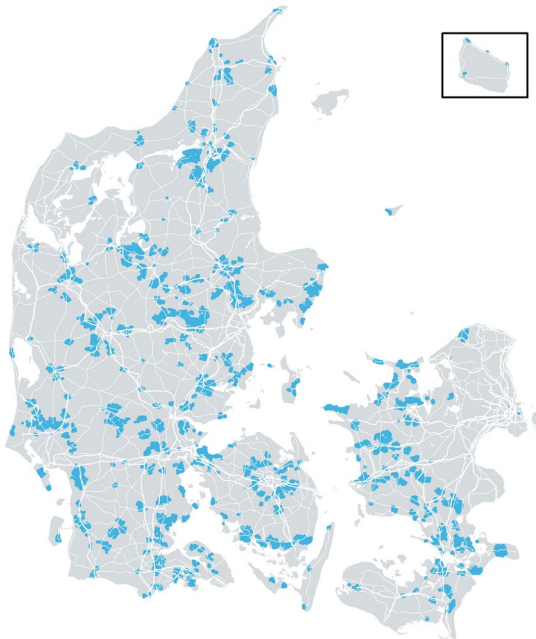
a) 1992



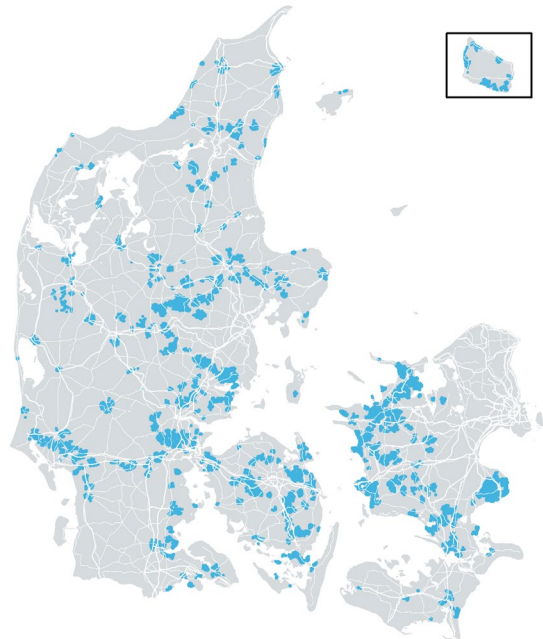
b) 2021



c) Forsvundne områder i 2021



d) Nye områder i 2021



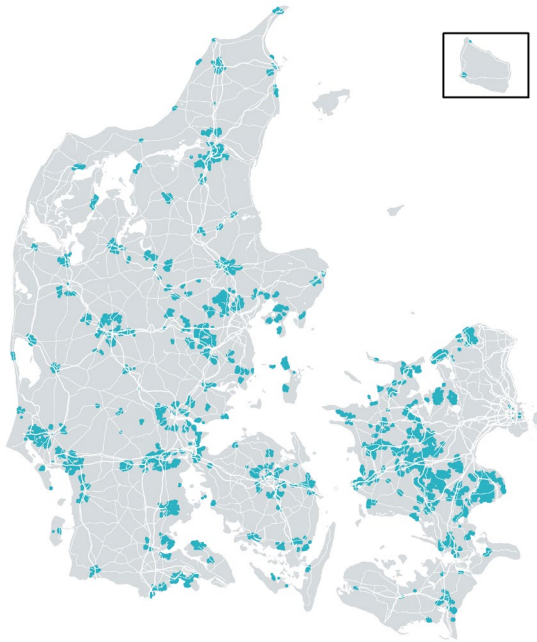
Anm.: Alle ejerboliger er rangeret efter boligens skønnede kvadratmeterpris. Herefter inddeles i ti lige store grupper (deciler) fra 1. decilgruppe indeholder den tiendedel af boligerne med den laveste kvadratmeterpris til 10. decil med de ti pct. af boligerne med den største kvadratmeterpris. Figuren viser placeringen af boliger i 3. decil i både 1992 og 2021. Desuden vises boliger som forlader og tilkommer gruppen i perioden 1992-2021.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

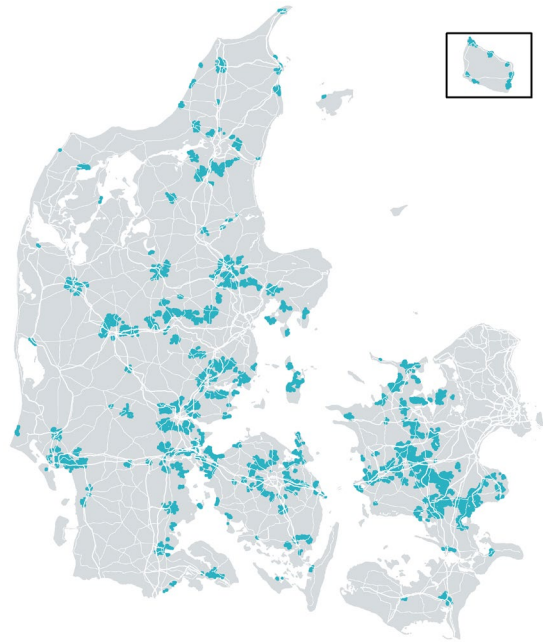
**Figur A.4**

**Ejerboliger med kvadratmeterpris i 4. decil**

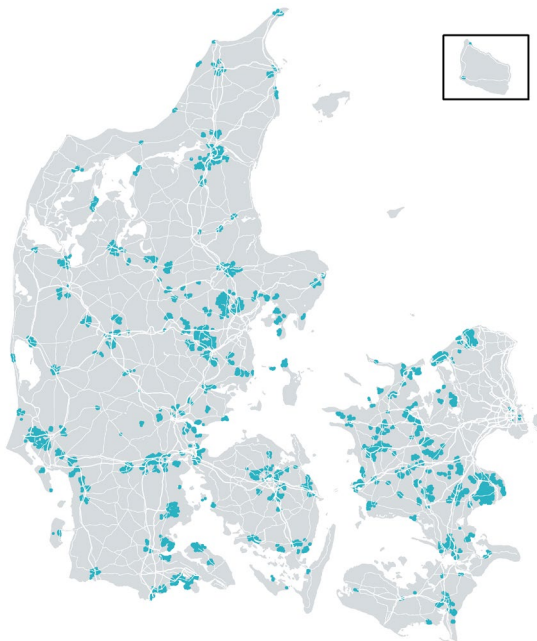
a) 1992



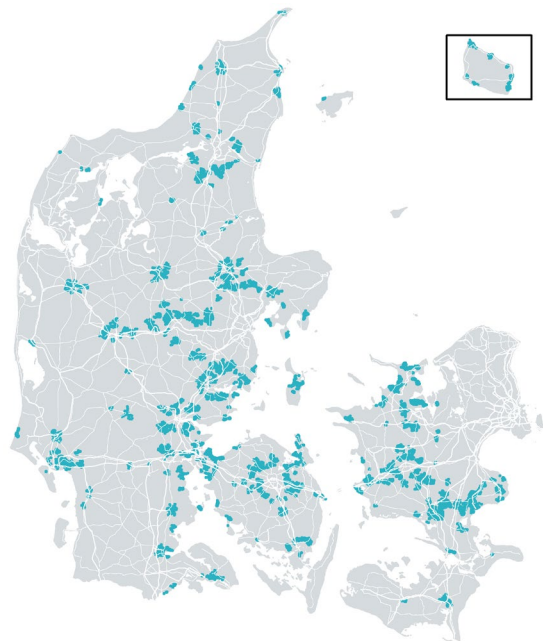
b) 2021



c) Forsvundne områder i 2021



d) Nye områder i 2021



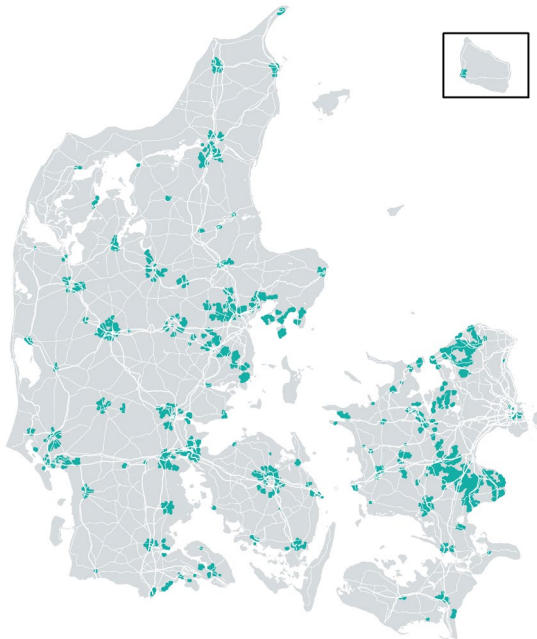
Anm.: Alle ejerboliger er rangeret efter boligens skønnede kvadratmeterpris. Herefter inddeles i ti lige store grupper (deciler) fra 1. decilgruppe indeholder den tiendedel af boligerne med den laveste kvadratmeterpris til 10. decil med de ti pct. af boligerne med den største kvadratmeterpris. Figuren viser placeringen af boliger i 4. decil i både 1992 og 2021. Desuden vises boliger som forlader og tilkommer gruppen i perioden 1992-2021.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

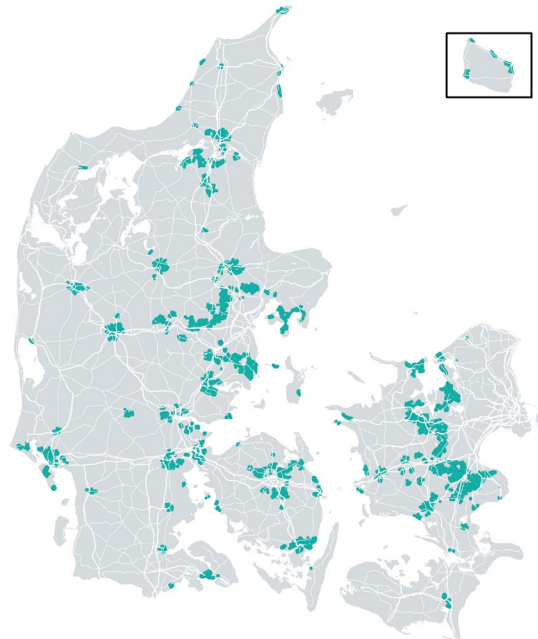
**Figur A.5**

**Ejerboliger med kvadratmeterpris i 5. decil**

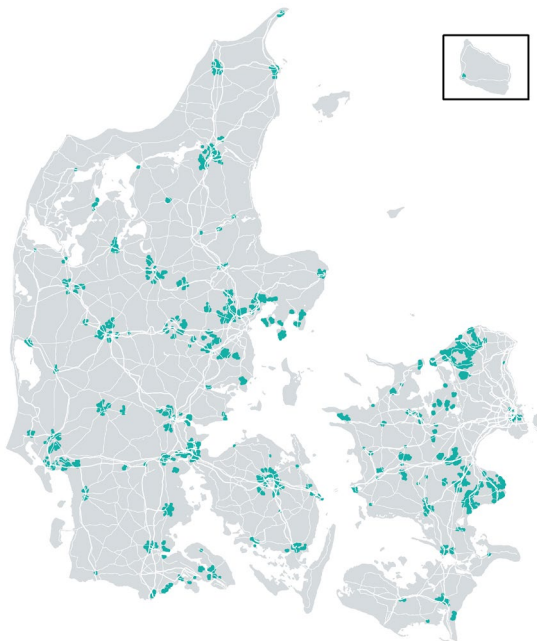
a) 1992



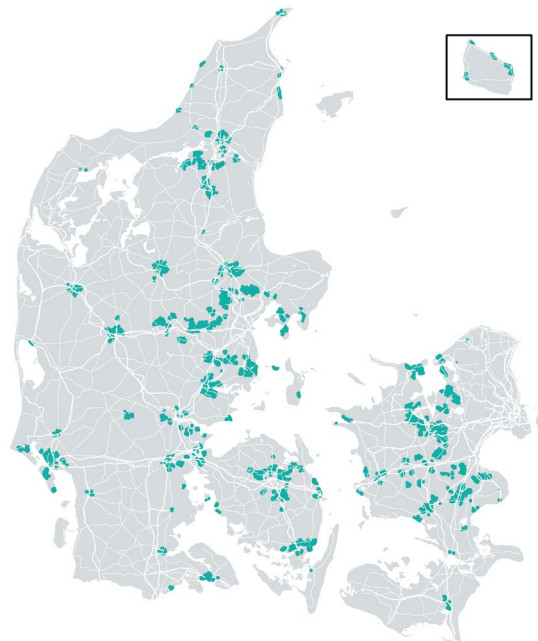
b) 2021



c) Forsvundne områder i 2021



d) Nye områder i 2021



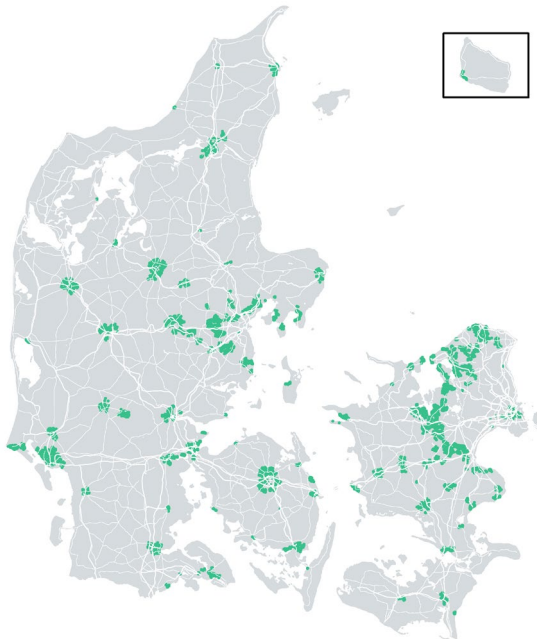
Anm.: Alle ejerboliger er rangeret efter boligens skønnede kvadratmeterpris. Herefter inddeles i ti lige store grupper (deciler) fra 1. decilgruppe indeholder den tiendedel af boligerne med den laveste kvadratmeterpris til 10. decil med de ti pct. af boligerne med den største kvadratmeterpris. Figuren viser placeringen af boliger i 5. decil i både 1992 og 2021. Desuden vises boliger som forlader og tilkommer gruppen i perioden 1992-2021.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

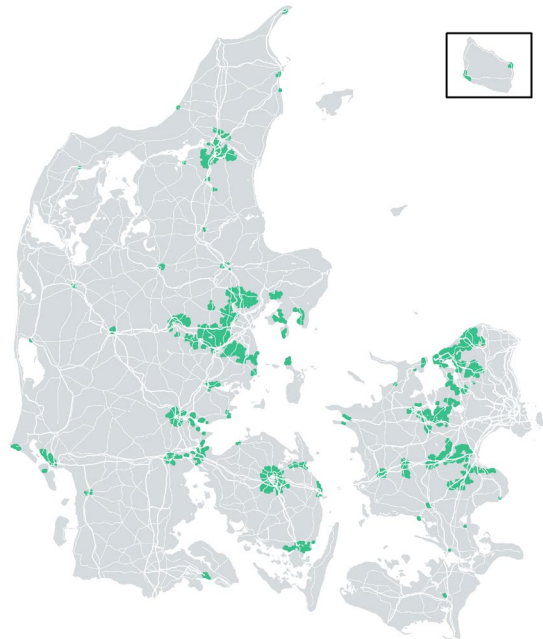
Figur A.6

Ejerboliger med kvadratmeterpris i 6. decil

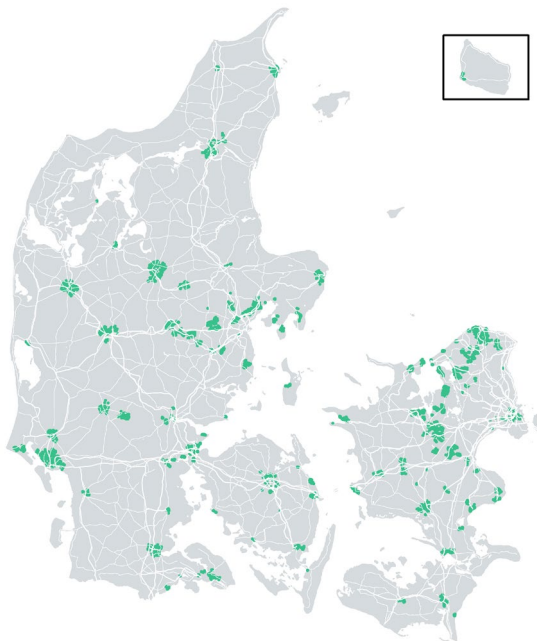
a) 1992



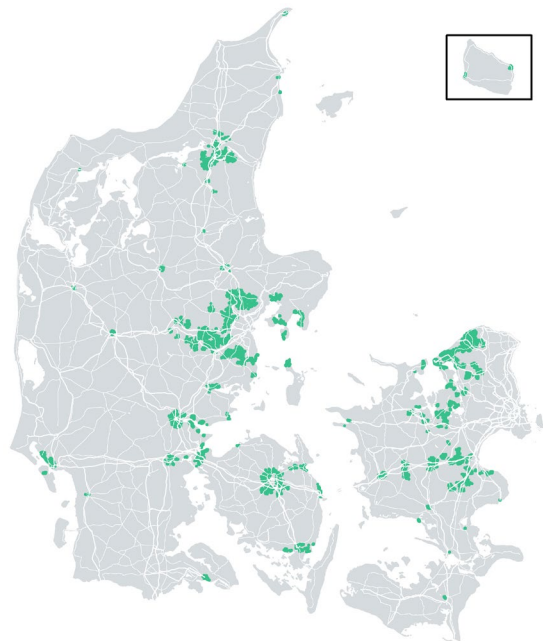
b) 2021



c) Forsvundne områder i 2021



d) Nye områder i 2021



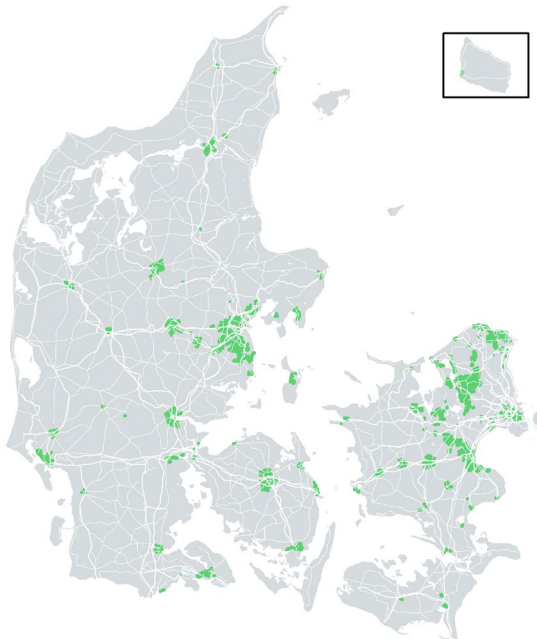
Anm.: Alle ejerboliger er rangeret efter boligens skønnede kvadratmeterpris. Herefter inddeles i ti lige store grupper (deciler) fra 1. decilgruppe indeholder den tiendedel af boligerne med den laveste kvadratmeterpris til 10. decil med de ti pct. af boligerne med den største kvadratmeterpris. Figuren viser placeringen af boliger i 6. decil i både 1992 og 2021. Desuden vises boliger som forlader og tilkommer gruppen i perioden 1992-2021.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

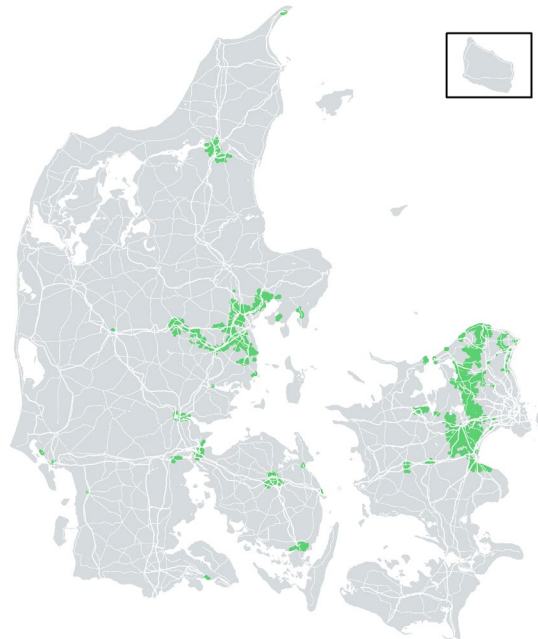
Figur A.7

Ejerboliger med kvadratmeterpris 7.decil

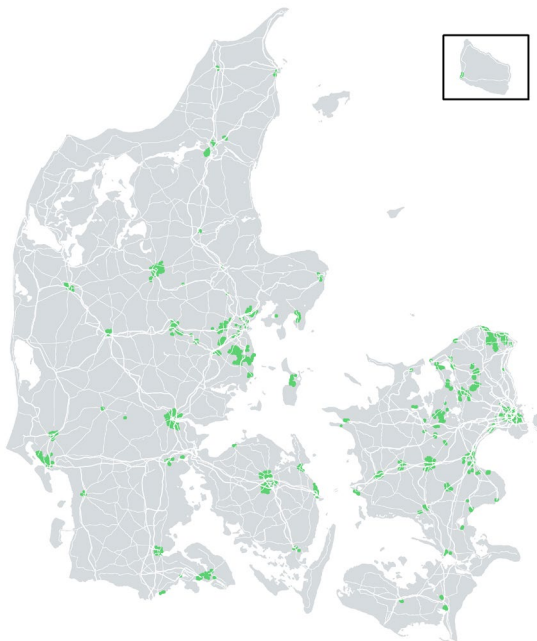
a) 1992



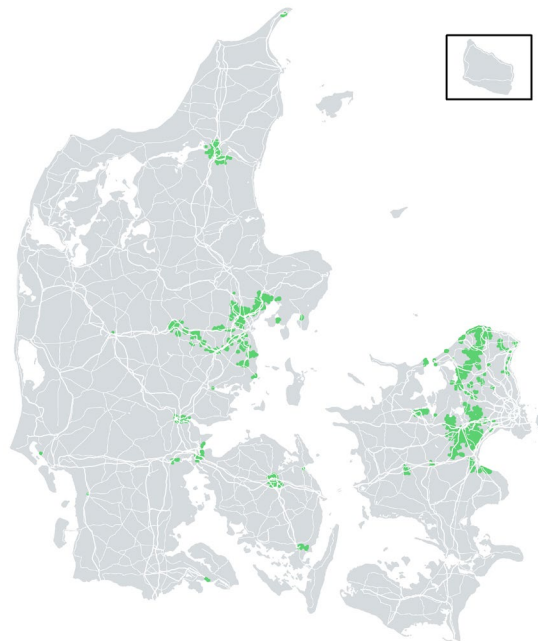
b) 2021



c) Forsvundne områder i 2021



d) Nye områder i 2021



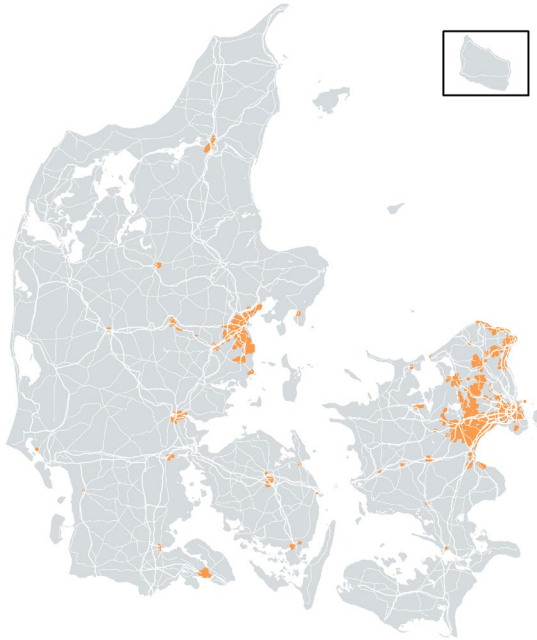
Anm.: Alle ejerboliger er rangeret efter boligens skønnede kvadratmeterpris. Herefter inddeles i ti lige store grupper (deciler) fra 1. decilgruppe indeholder den tiendedel af boligerne med den laveste kvadratmeterpris til 10. decil med de ti pct. af boligerne med den største kvadratmeterpris. Figuren viser placeringen af boliger i 7. decil i både 1992 og 2021. Desuden vises boliger som forlader og tilkommer gruppen i perioden 1992-2021.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur A.8

Ejerboliger med kvadratmeterpris i 8. decil

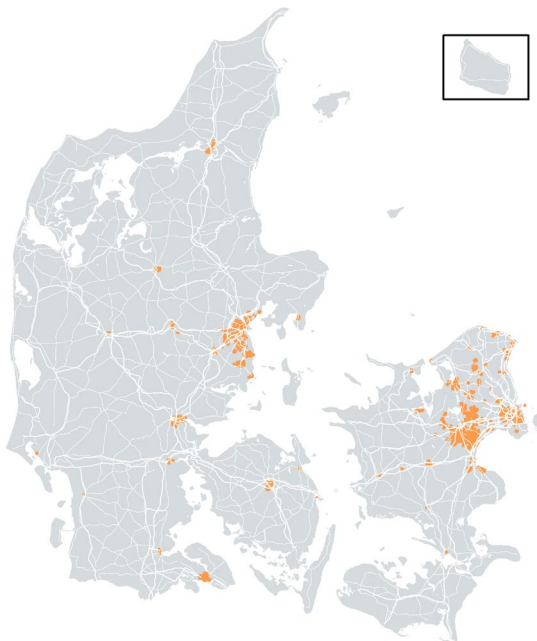
a) 1992



b) 2021



c) Forsvundne områder i 2021



d) Nye områder i 2021



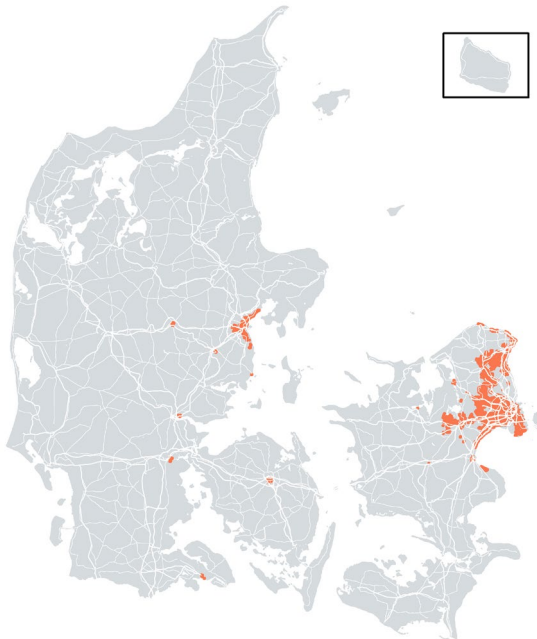
Anm.: Alle ejerboliger er rangeret efter boligens skønnede kvadratmeterpris. Herefter inddeles i ti lige store grupper (deciler) fra 1. decilgruppe indeholder den tiendedel af boligerne med den laveste kvadratmeterpris til 10. decil med de ti pct. af boligerne med den største kvadratmeterpris. Figuren viser placeringen af boliger i 8. decil i både 1992 og 2021. Desuden vises boliger som forlader og tilkommer gruppen i perioden 1992-2021.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur A.9

Ejerboliger med kvadratmeterpris i 9. decil

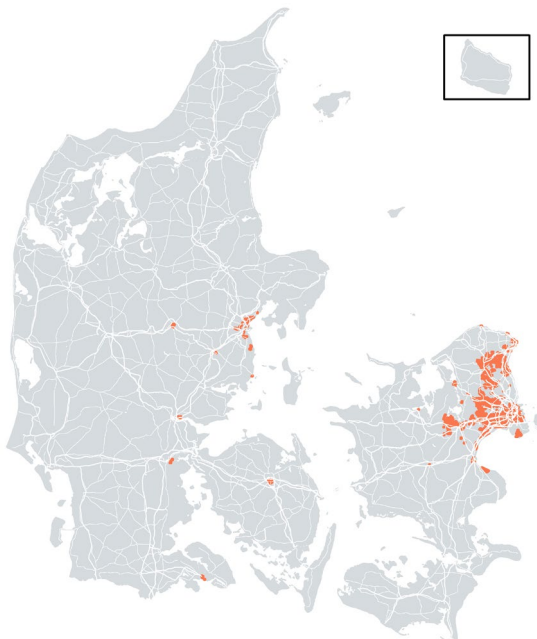
a) 1992



b) 2021



c) Forsvundne områder i 2021



d) Nye områder i 2021



Anm.: Alle ejerboliger er rangeret efter boligens skønnede kvadratmeterpris. Herefter inddeles i ti lige store grupper (deciler) fra 1. decilgruppe indeholder den tiendedel af boligerne med den laveste kvadratmeterpris til 10. decil med de ti pct. af boligerne med den største kvadratmeterpris. Figuren viser placeringen af boliger i 9. decil i både 1992 og 2021. Desuden vises boliger som forlader og tilkommer gruppen i perioden 1992-2021.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

Figur A.10

Ejerboliger med kvadratmeterpris i 10. decil

a) 1992



b) 2021



c) Forsvundne områder i 2021



d) Nye områder i 2021



Anm.: Alle ejerboliger er rangeret efter boligens skønnede kvadratmeterpris. Herefter inddeles i ti lige store grupper (deciler) fra 1. decilgruppe indeholder den tiendedel af boligerne med den laveste kvadratmeterpris til 10. decil med de ti pct. af boligerne med den største kvadratmeterpris. Figuren viser placeringen af boliger i 10. decil i både 1992 og 2021. Desuden vises boliger som forlader og tilkommer gruppen i perioden 1992-2021.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.



## B. Årlig vækstrate i boligpriserne efter region

I det følgende fokuseres på den geografiske udvikling i ejerboligpriserne for de to delperioder, hvor priserne steg mest markant. Dette er årene op til finanskrisen, afgrænset som perioden 1999-2006, og årene 2013-2021, hvor økonomien igen var kommet sig over krisen.

Ejerboligprisudviklingen for de to opsvingsperioder beskrives, sammenholdes og illustreres ud fra en række landkort. Kortene viser i figur a) og c) den gennemsnitlige årlige procentuelle vækst i ejerboligens kvadratmeterpris. Herudover viser figur b) og d) den gennemsnitlige årlige vækst i 2021-kr.

Figur 4.7 på side 49 viser den geografiske ejerboligprisudvikling for hele landet. I dette appendiks vises tilsvarende landkort for udvalgte landsdele og byområder.

### Hele landet

Over hele landet var prisudviklingen for ejerboliger positiv i begge opsvingsperioder. Den gennemsnitlige årlige stigningstakt varierede imidlertid betydeligt afhængig af boligens geografiske placering.

Prisstigningen var i begge perioder mest markant omkring de større byer og mindre udtalt i landdistrikterne. I perioden 1999-2006 steg priserne markant mere end de gjorde i perioden 2013-2021. Og i den første opgangsperiode dækkede prisstigningerne et større geografisk område. I perioden 2013-2021 var prisstigningerne mere lokalt forankret i og tæt omkring de større byer.

På landsplan varierer den årlige stigningstakt i kvadratmeterprisen for en ejerbolig fra under 2 pct. til over 8 pct. afhængig af boligens geografiske placering. Den høje prisvækst havde større geografiske udbredelse i perioden 1999-2006, end det var tilfældet i perioden 2013-2021. På stort set hele Sjælland og omkring de største byer – særligt i og omkring Aarhus – steg priserne i årene 1999-2006 med over 8 pct. årligt. Dette var ligeledes tilfældet for dele af Lolland, Falster og Sydfyn.

I Aarhus, København og Nordsjælland svarede væksten til en gennemsnitlig stigning i kvadratmeterpriserne på over 2.000 kr. pr. år. Mens det på Lolland, Falster og Sydfyn svarede til en gennemsnitlig årlig stigning på omkring 250-500 kr. pr. kvadratmeter.

I perioden efter 2013 var det færre områder, der oplevede samme vækst i ejerboligpriserne som op til finanskrisen. Dette var eksempelvis omkring Hovedstaden og på Bornholm. For områderne ved Aarhus, Nordsjælland og Sydfyn var prisstigningerne i denne periode på under 4 pct., hvilket også gjorde sig gældende for Lolland og Falster.

Ejerboligerne i Nordsjælland starter her ved en relativ høj kvadratmeterpris i 2013, så prisstigningen på omkring 4 pct. svarede til en gennemsnitlig årlig stigning på omkring 2.000 kr. pr. kvadratmeter. Altså samme nominelle niveau som perioden op til finanskrisen. Ejerboligerne på eksempelvis Lolland og Falster kom fra et relativt lavt prisniveau, hvorved væksten her svarede til en gennemsnitlig årlig stigning i kvadratmeterpriserne på omkring 100 kr. Dette pointerer vigtigheden af at tage højde for boligprisens udgangspunkt (baseline-effekten), når de geografiske vækstrater vurderes.

## Udvalgte geografiske områder

### Københavnsområdet

I det indre København og i Storkøbenhavn var prisudviklingen generelt ensartet geografisk fordelt i begge opgangsperioder 1999-2006 og 2013-2021, jf. Figur B.1 og Figur B.2. Således var det omtrent de samme boliger, der oplevede høj vækst i priserne i begge tidsperioder. Andre ejerboliger oplevede lavere vækst i begge perioder.

Det fremgår, at den første af opsvingsperiode 1999-2006 var kraftigere end den sidste 2013-2021. Når den årlige procentuelle vækst i boligens kvadratmeterpris betragtes er det værd at huske den ovenover beskrevne prisniveau baseline-effekt. Op til finanskrisen var den gennemsnitlige årlige prisstigning på over 12 pct. for hovedparten af ejerboliger i indre København. Dette svarede til en ændring i kvadratmeterpriserne på omkring 3.500 kr. pr. år. I perioden 2013-2021 var stigningerne ikke lige så markante. De fleste boliger steg med mellem 6 og 8 pct. om året. Dette svarede dog til en ændring i kvadratmeterpriserne pr. år på stort set samme niveau som den første opsvingsperiode.

### Sjælland

Hvis hele Sjælland betragtes ses en større variation i den årlige stigningstakt i ejerboligprisen pr. kvadratmeter, end det var tilfældet i Københavnsområdet, jf. Figur B.3. Generelt set er væksten i priserne lavere, jo længere væk man bevæger sig fra det indre København. Dog ses de større byer uden for København at være lokale centre med forholdsvis høj årlig vækst i boligpriserne. Dette gælder for eksempel områderne omkring Slagelse- Sorø, Næstved og Nykøbing Falster.

Væksten i priserne har været højere og udbredt over et større geografisk område i perioden 1999-2006 end i perioden 2013-2021. Forskellen mellem perioderne er igen til dels forklaret af forskellige prisniveauer i udgangspunktet.

I modsætning til tendensen i København og Storkøbenhavn kan man på det øvrige Sjælland ikke i samme grad observere den samme geografiske udbredelse i prisstigningerne mellem de to perioder.

For eksempel ses det, at der har været en relativ høj prisstigning i byerne langs motorvejen fra København og vestover i perioden 1999-2006, hvilket for eksempel gælder for byerne Ringsted, Sorø og Slagelse. Denne tendens ses ikke i perioden 2013-2021. Samme tendens gør sig gældende længere sydpå i de kystnære områder Skælskør, Bisserup og Karrebæksminde samt Næstved.

### Lolland og Falster

På Lolland og Falster har prisstigningerne været forskellige i de to perioder, men har dog været rimelig ensartet geografisk fordelt, jf. Figur B.4. Nykøbing Falster og Marielyst skiller sig ud, da ejerboligerne her i perioden 1999-2006 havde højere nominelle prisstigninger relativt til de øvrige områder. Denne tendens ses ikke i perioden 2013-2021.

Sommerhushandler indgår ikke som en del af analysen, hvorfor de ikke burde have en indvirkning på udviklingen i eksempelvis Marielyst, hvor sommerhuse udgør en stor del af ejendommene. Når det alligevel observeres, at prisstigningerne i områder med mange sommerhuse er højere end andre tilsvarende områder, kan det tænkes, at enkelte sommerhuse er registreret forkert, og dermed alligevel er blevet medtaget i analysen.

### Landsdel Fyn

Figur B.5 viser den samme udvikling på Fyn som i resten af landet. Prisstigningerne blev mindre jo længere man bevæger sig væk fra Odense, og kystbyerne. I årene op til finanskrisen

oplevedes høje prisstigninger i Odense, og langs hele Fyns kystlinje. I byer som Bogense, Assens og Fåborg steg priserne tilnærmelsesvis i samme grad som i centrum af Odense. I Middelfart, Svendborg og Odense svarede det til en stigning i kvadratmeterpriser på over 1.000 kr. årligt. Mellem 2013 og 2021 var det kun i Odense, at priserne forsat steg markant. Her, og i Svendborg og Middelfart, svarede dette fortsat til en stigning på over 1.000 kr. årligt.

### **Landsdel Syddjylland**

I Syddjylland havde et større område omkring Kolding- og Vejlefjord oplevet relativt store prisstigninger i den første af opsvingsperioderne 1999-2006, jf. Figur B.6. I perioden 2013-2021 var området indsnævret til, at det nu kun var selve Vejle og Kolding, der havde oplevet større prisstigninger, men nu i et mere afdæmpet omfang.

Ejerboliger i Jelling og Grønninghoved Strand oplevede i perioden 1999-2006 også markante prisstigninger. Dette gjaldt ligeledes områder længere sydpå, eksempelvis Aabæk tæt på Åbenrå samt Sønderborg og Rønshoved. Også byer på Vestkysten som Ribe, Esbjerg og Blåvand oplevede stigninger på over 7 pct. årligt i perioden 1999-2006, hvilket her svarede til en stigning i kvadratmeterprisen på over 700 kr. årligt. I disse områder var prisstigningerne ikke i samme størrelsesorden i perioden 2013-2021, og der ses generelt relativt mindre prisstigninger for hele Syddjylland i denne periode.

### **Region Midtjylland**

Også i region Midtjylland er det tydeligt, at prisstigningerne 1999-2006 omfattede et langt større område, end det var tilfældet i årene 2013-2021, jf. Figur B.7. Et stort område omkring Aarhus som strækker sig fra Ebeltoft, nordpå til Randers og hele vejen til Silkeborg, Skanderborg og Horsens oplevede prisstigninger på gennemsnitligt over 8 pct. om året i perioden 1999-2006. I Aarhus var der tale om en stigning i kvadratmeterprisen på over 2.000 kr. årligt. Fra 2013-2021 ses et mere indskrænket og mere afdæmpet højvækstområde. I Aarhus steg kvadratmeterpriserne dog også i perioden 2013-2021 med over 2.000 kr. årligt.

### **Region Nordjylland**

I region Nordjylland ses den kraftigste årlige vækst ejerboligprisen i Aalborg, jf. Figur B.8 og Figur B.9. Dette gælder begge opsvingsperioder 1999-2006 og 2013-2021.

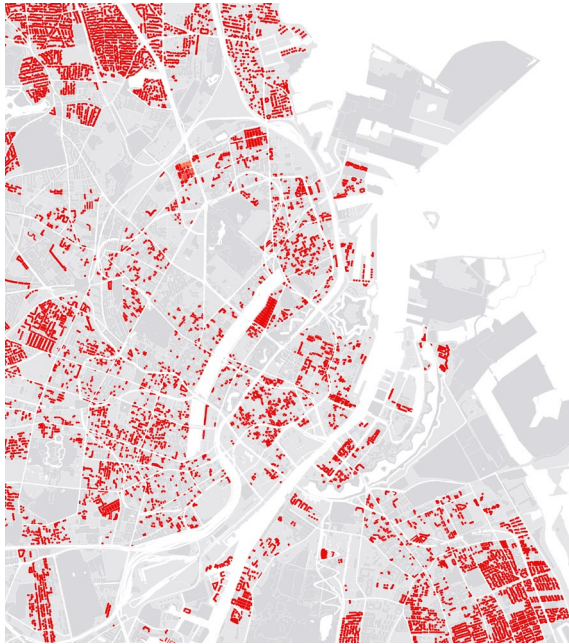
Især bydelen Hasseri og det vestlige Nørresundby havde i begge perioder oplevet høje prisstigninger. I perioden 1999-2006 var den gennemsnitlige årlige vækst over 7 pct. i disse bydele, hvilket svarede til at kvadratmeterpriserne i denne periode steg med over 1.000 kr. om året. I perioden 2013-2021 var den gennemsnitlige årlige vækst omkring 4 pct. i områderne, hvilket tilnærmelsesvis var den samme som for hele Aalborg. For Hasseri svarede væksten i denne periode til en årlig stigning i priserne på 900-1.000 kr. Ejerboliger beliggende i Aalborg SV og Aalborg Øst havde i perioderne oplevet en lavere prisstigning i forhold til resten af Aalborg.

Ud over Aalborg oplevede en række andre områder i Nordjylland også høje prisstigninger i perioden 1999-2006. Herunder Hobro, dele af Læsø samt kystbyer som Blokhus, Løkken og Skagen. Denne tendens er ikke så tydelig i den efterfølgende opsvingsperiode. Især Hobro oplevede i perioden 2013-2021 en relativt svagere stigning i priserne.

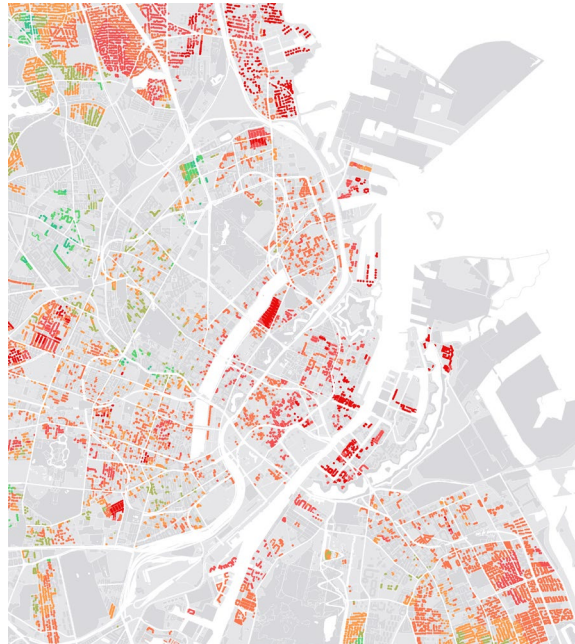
**Figur B.1**

Gennemsnitlig årlig ændring i boligprisen pr. kvadratmeter, København, 1999-2006 og 2013-2021

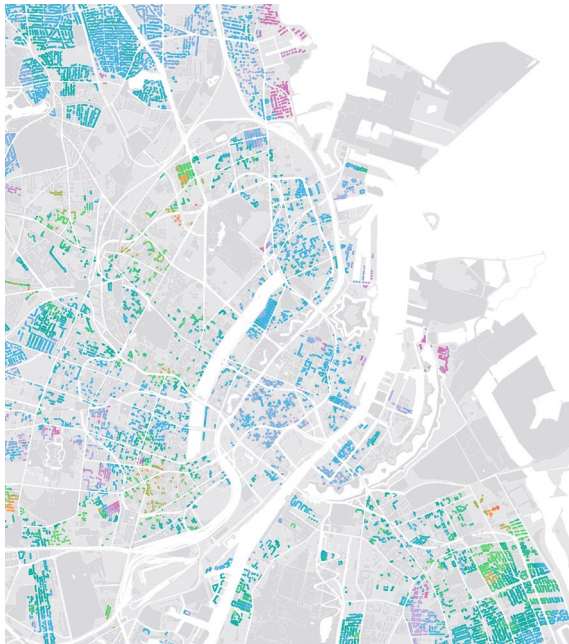
a) Gns. årlig ændring 1999-2006, pct.



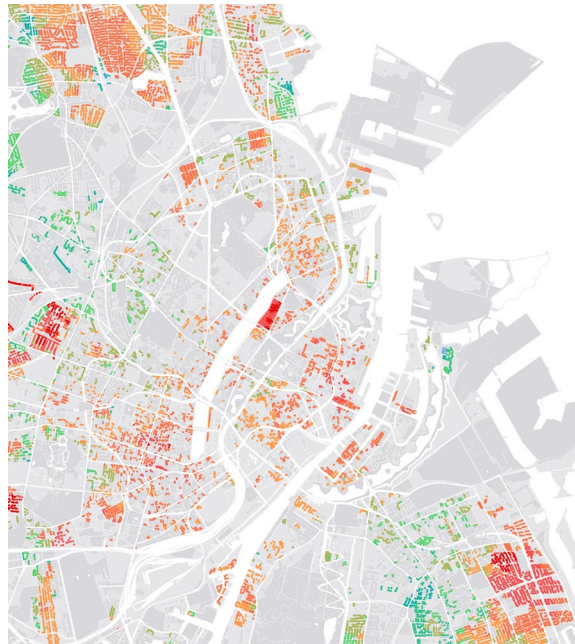
b) Gns. årlig ændring 1999-2006, kr. (2021-niveau)



c) Gns. årlig ændring 2013-2021, pct.



d) Gns. årlig ændring 2013-2021, kr. (2021-niveau)



< 6 pct.	6-6,5	6,5-7	7-7,5
7,5-8	8-8,5	8,5-9	9-9,5
9,5-10	10-10,5	10,5-11	11-11,5
11,5-12	> 12 pct.		

< 1.100 kr.	1.100-1.200	1.200-1.300	1.300-1.400
1.400-1.500	1.500-1.750	1.750-2.000	2.000-2.250
2.250-2.500	2.500-2.750	2.750-3.000	3.000-3.500
3.500-4.000	> 4.000 kr.		

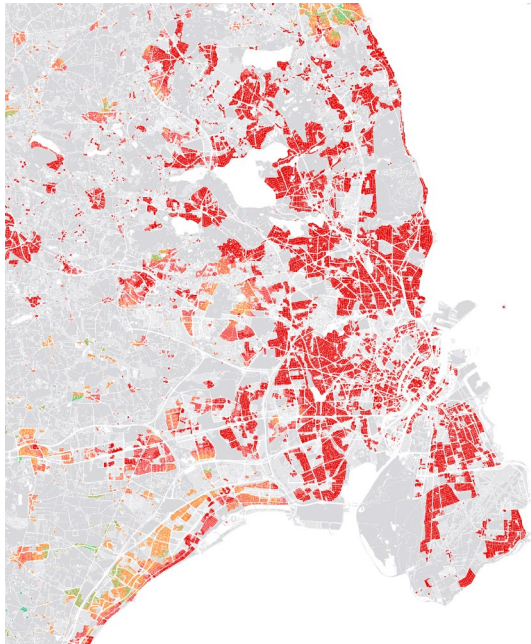
Anm.: Figuren viser den estimerede gennemsnitlige årlige ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

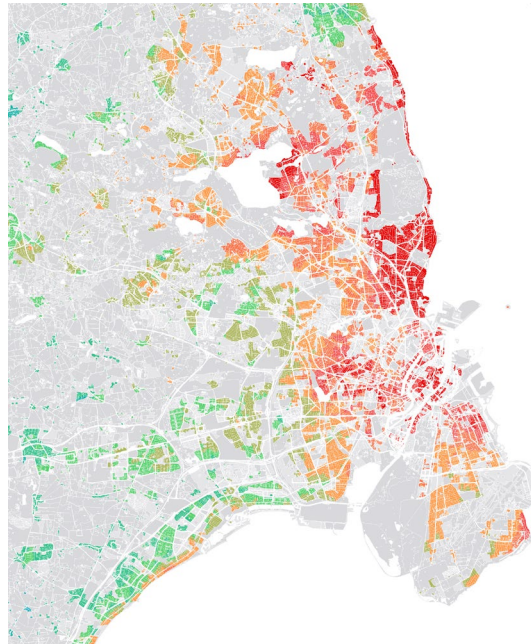
**Figur B.2**

Gennemsnitlig årlig ændring i boligprisen pr. kvadratmeter, Storkøbenhavn, 1999-2006 og 2013-2021

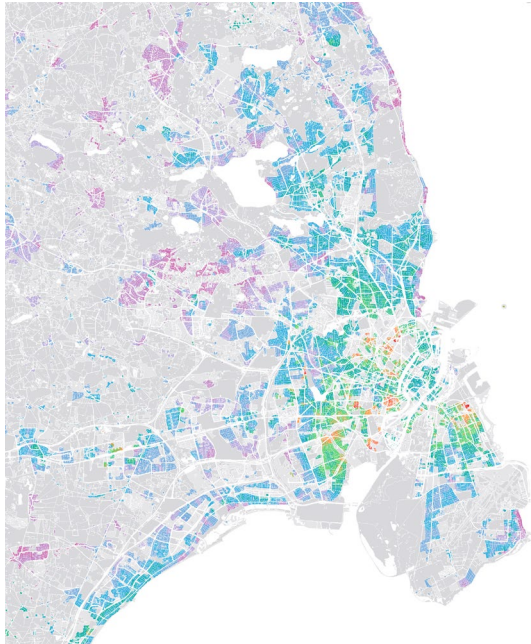
a) Gns. årlig ændring 1999-2006, pct.



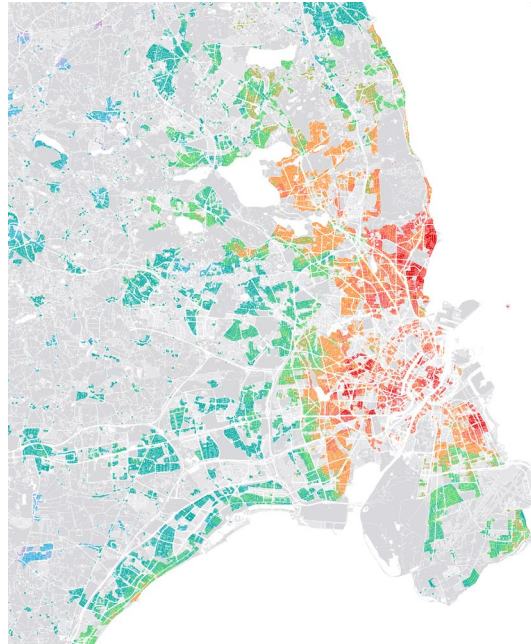
b) Gns. årlig ændring 1999-2006, kr. (2021-niveau)



c) Gns. årlig ændring 2013-2021, pct.



d) Gns. årlig ændring 2013-2021, kr. (2021-niveau)



< 5 pct.	5-5,5	5,5-6	6-6,5	< 500 kr.	500-600	600-700	700-800
6,5-7	7-7,5	7,5-8	8-8,5	800-900	900-1.000	1.000-1.250	1.250-1.500
8,5-9	9-9,5	9,5-10	10-10,5	1.500-1.750	1.750-2.000	2.000-2.500	2.500-3.000
10,5-11	> 11 pct.			3.000-3.500	> 3.500 kr.		

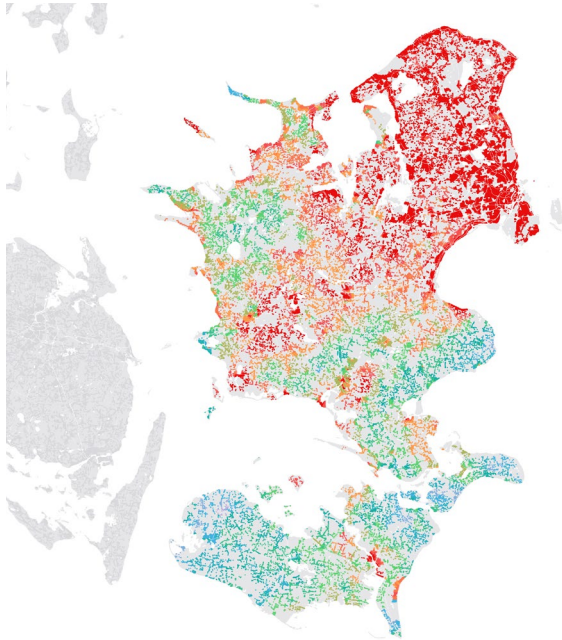
Anm.: Figuren viser den estimerede gennemsnitlige årlige ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

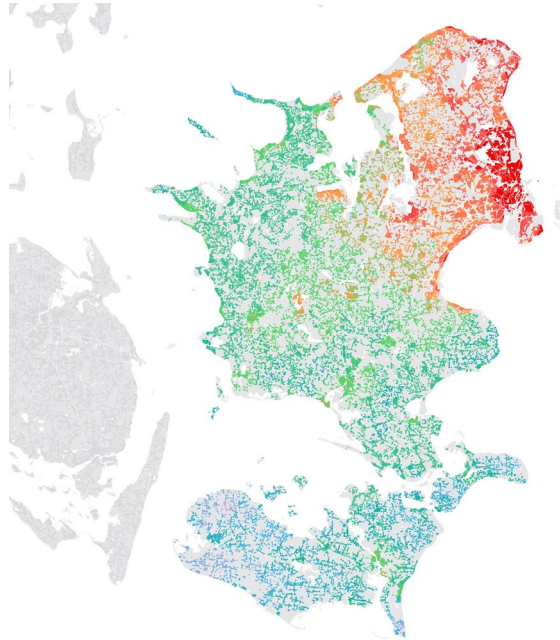
**Figur B.3**

Gns. årlig ændring i boligprisen pr. m<sup>2</sup>, region Hovedstaden og Sjælland, 1999-2006 og 2013-2021

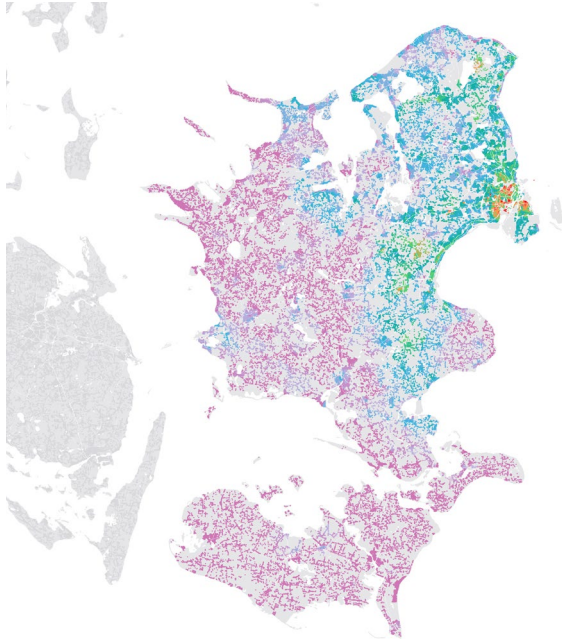
a) Gns. årlig ændring 1999-2006, pct.



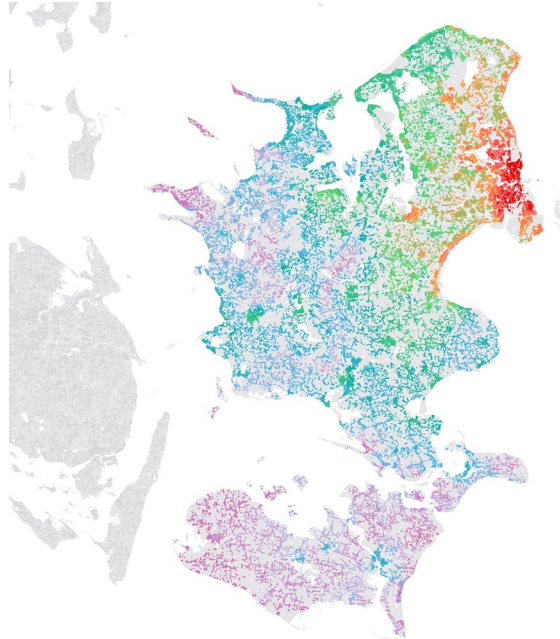
b) Gns. årlig ændring 1999-2006, kr. (2021-niveau)



c) Gns. årlig ændring 2013-2021, pct.



d) Gns. årlig ændring 2013-2021, kr. (2021-niveau)



< 4 pct.	4-4,5	4,5-5	5-5,5	< 100 kr.	100-150	150-200	200-250
5,5-6	6-6,5	6,5-7	7-7,5	250-300	300-400	400-500	500-750
7,5-8	8-8,5	8,5-9	9-9,5	750-1.000	1.000-1.250	1.250-1.500	1.500-2.000
9,5-10	> 10 pct.			2.000-2.500	> 2.500 kr.		

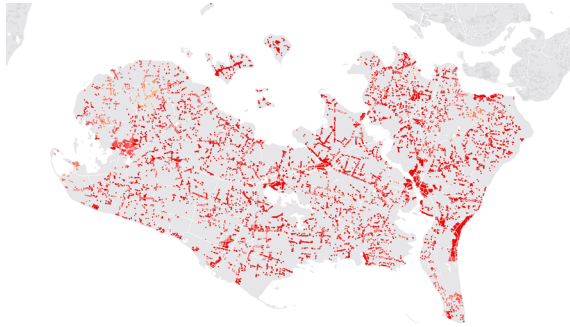
Anm.: Figuren viser den estimerede gennemsnitlige årlige ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

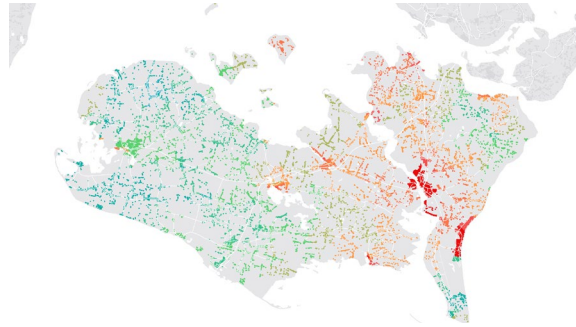
**Figur B.4**

Gennemsnitlig årlig ændring i boligprisen pr. kvadratmeter, Lolland og Falster, 1999-2006 og 2013-2021

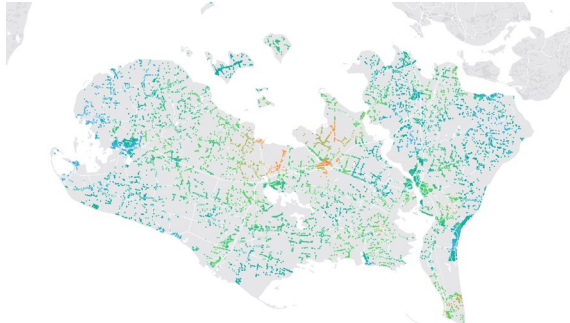
a) Gns. årlig ændring 1999-2006, pct.



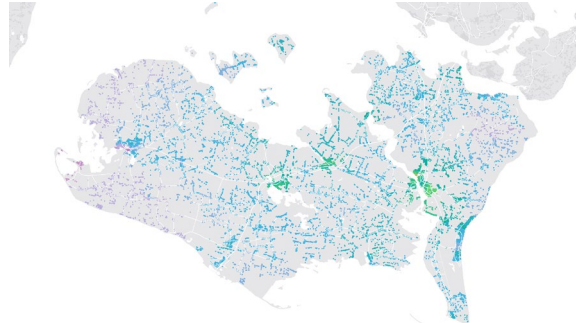
b) Gns. årlig ændring 1999-2006, kr. (2021-niveau)



c) Gns. årlig ændring 2013-2021, pct.



d) Gns. årlig ændring 2013-2021, kr. (2021-niveau)



< 0 pct.	0-0,5	0,5-1	1-1,5	< 20 kr.	20-50	50-80	80-100
1,5-2	2-2,5	2,5-3	3-3,5	100-150	150-200	200-250	250-300
3,5-4	4-4,5	4,5-5	5-5,5	200-350	350-400	400-500	500-600
5,5-6	> 6 pct.			600-700	> 700 kr.		

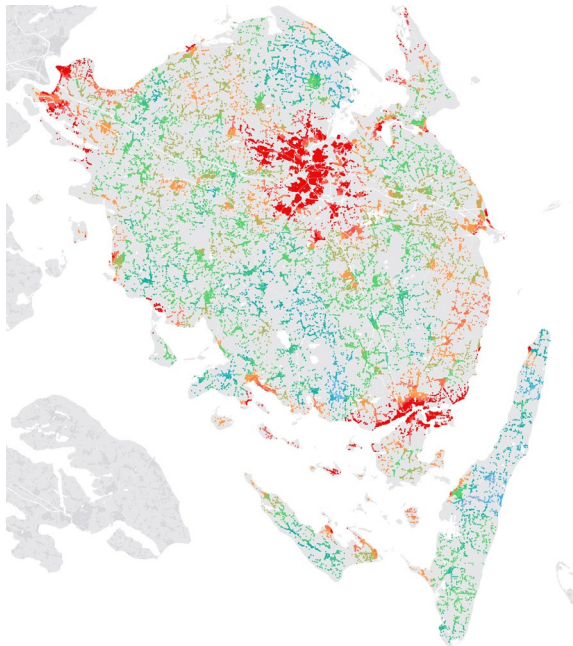
Anm.: Figuren viser den estimerede gennemsnitlige årlige ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

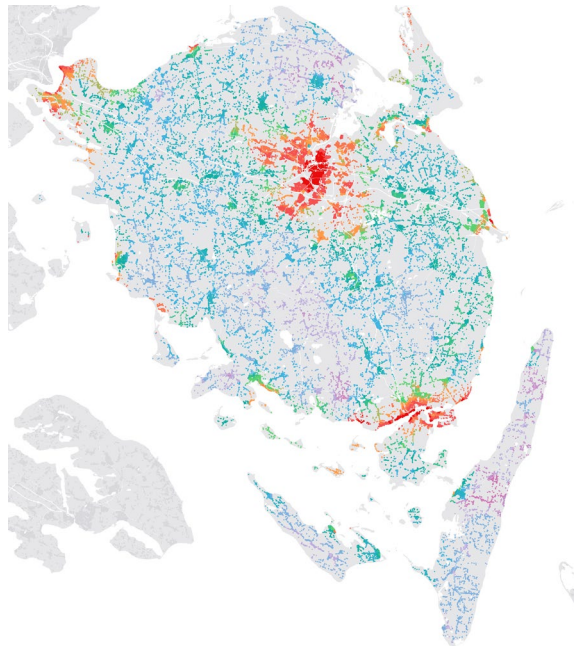
**Figur B.5**

Gennemsnitlig årlig ændring i boligprisen pr. kvadratmeter, landsdel Fyn, 1999-2006 og 2013-2021

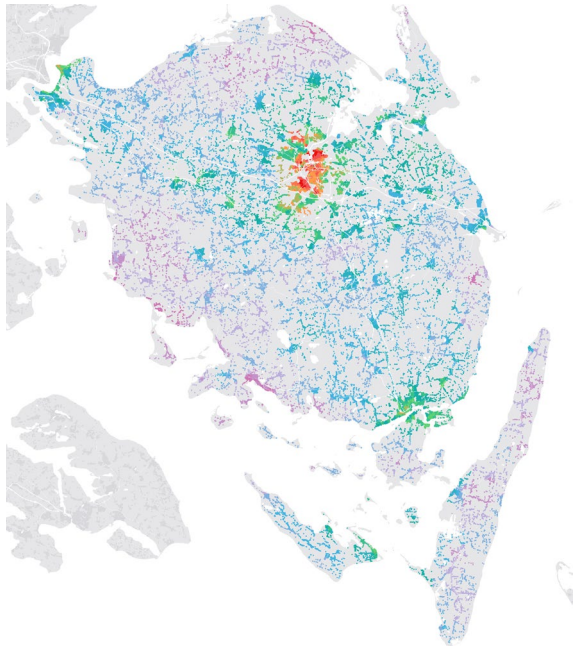
a) Gns. årlig ændring 1999-2006, pct.



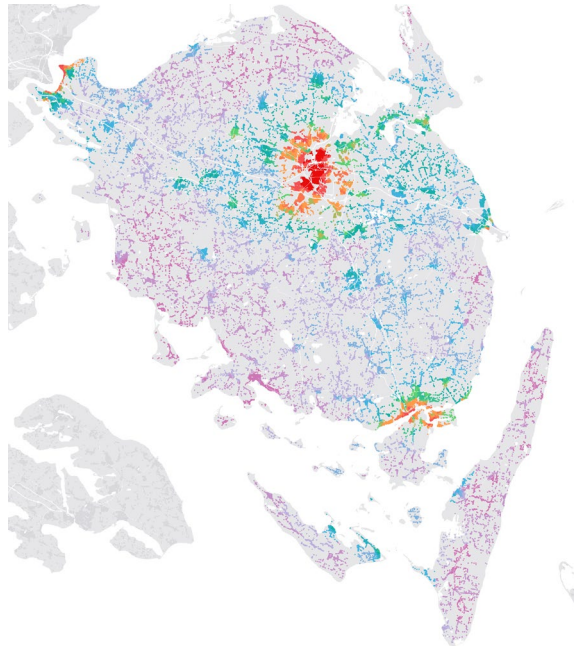
b) Gns. årlig ændring 1999-2006, kr. (2021-niveau)



c) Gns. årlig ændring 2013-2021, pct.



d) Gns. årlig ændring 2013-2021, kr. (2021-niveau)



< 1 pct.	1-1,5	1,5-2	2-2,5	< 50 kr.	50-100	100-150	150-200
2,5-3	3-3,5	3,5-4	4-4,5	200-250	250-300	300-350	350-400
4,5-5	5-5,5	5,5-6	6-6,5	400-450	450-500	500-600	600-700
6,5-7	> 7 pct.			700-1.000	> 1.000 kr.		

Anm.: Figuren viser den estimerede gennemsnitlige årlige ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

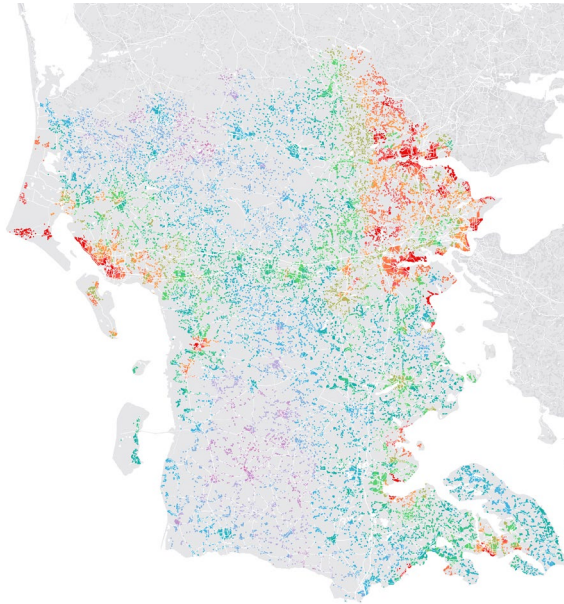
Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.



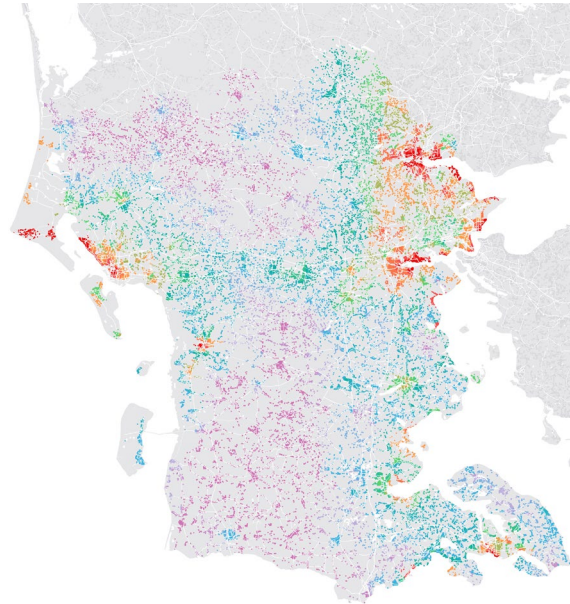
**Figur B.6**

Gennemsnitlig årlig ændring i boligprisen pr. kvadratmeter, landsdel Sydjylland, 1999-2006 og 2013-2021

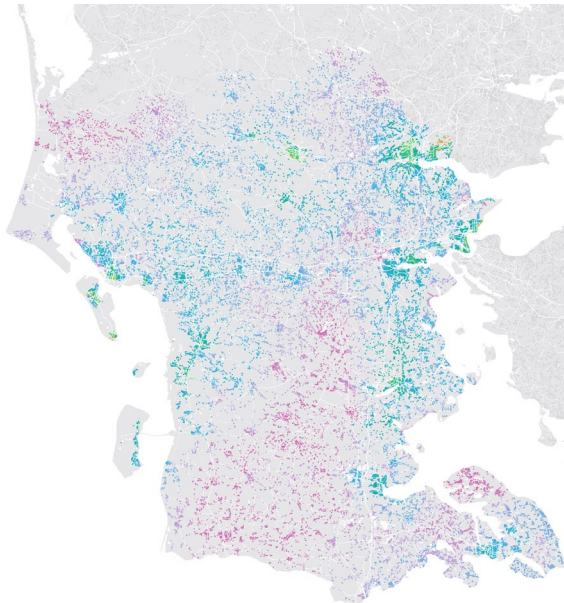
a) Gns. årlig ændring 1999-2006, pct.



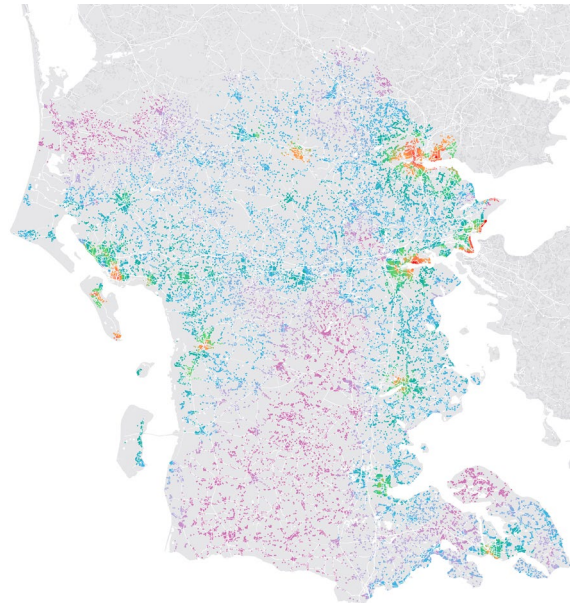
d) Gns. årlig ændring 1999-2006, kr. (2021-niveau)



c) Gns. årlig ændring 2013-2021, pct.



d) Gns. årlig ændring 2013-2021, kr. (2021-niveau)



< 1 pct.	1-1,5	1,5-2	2-2,5	< 20 kr.	20-50	50-80	80-100
2,5-3	3-3,5	3,5-4	4-4,5	100-150	150-200	200-250	250-300
4,5-5	5-5,5	5,5-6	6-6,5	200-350	350-400	400-500	500-600
6,5-7	> 7 pct.			600-700	> 700 kr.		

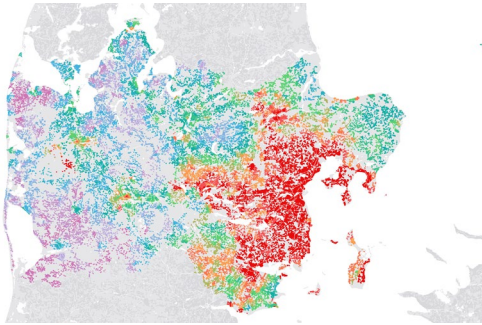
Anm.: Figuren viser den estimerede gennemsnitlige årlige ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

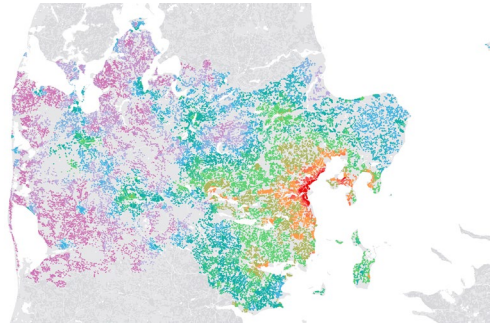
**Figur B.7**

Gennemsnitlig årlig ændring i boligprisen pr. kvadratmeter, region Midtjylland, 1999-2006 og 2013-2021

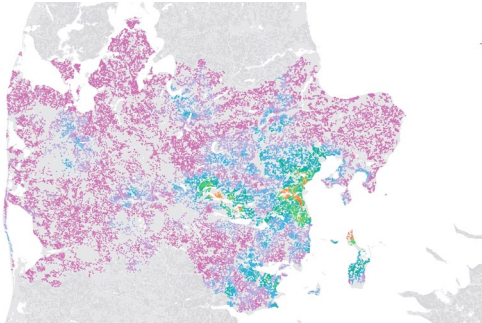
a) Gns. årlig ændring 1999-2006, pct.



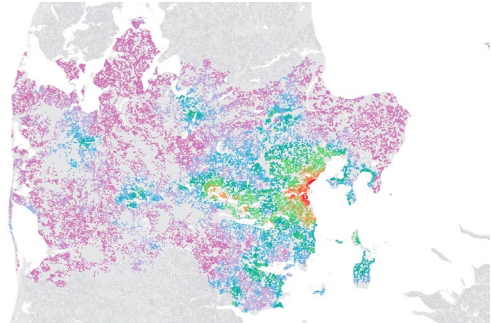
b) Gns. årlig ændring 1999-2006, kr. (2021-niveau)



c) Gns. årlig ændring 2013-2021, pct.



d) Gns. årlig ændring 2013-2021, kr. (2021-niveau)



< 2 pct.	2-2,5	2,5-3	3-3,5	< 50 kr.	50-100	100-150	150-200
3,5-4	4-4,5	4,5-5	5-5,5	200-250	250-300	300-400	400-500
5,5-6	6-6,5	6,5-7	7-7,5	500-750	750-1.000	1.000-1.250	1.250-1.500
7,5-8	> 8 pct.			1.500-2.000	> 2.000 kr.		

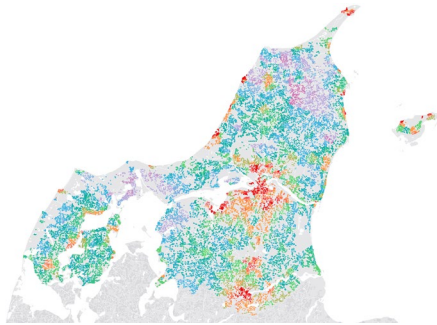
Anm.: Figuren viser den estimerede gennemsnitlige årlige ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

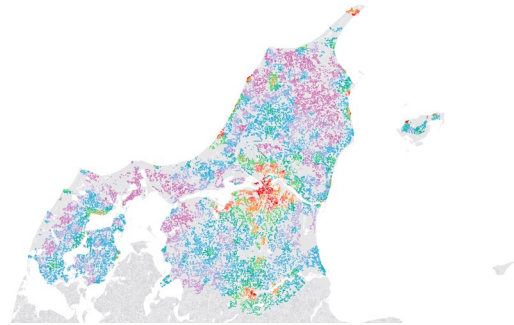
**Figur B.8**

Gennemsnitlig årlig ændring i boligprisen pr. kvadratmeter, region Nordjylland, 1999-2006 og 2013-2021

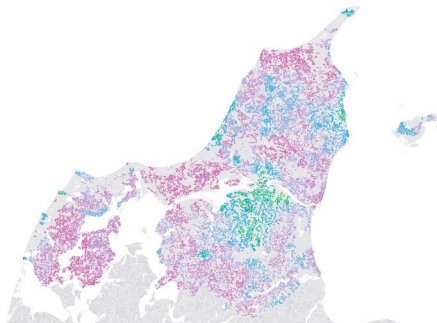
a) Gns. årlig ændring 1999-2006, pct.



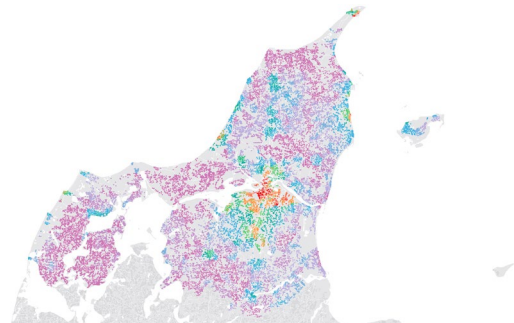
b) Gns. årlig ændring 1999-2006, kr. (2021-niveau)



c) Gns. årlig ændring 2013-2021, pct.



d) Gns. årlig ændring 2013-2021, kr. (2021-niveau)



< 1 pct.	1-1,5	1,5-2	2-2,5	< 20 kr.	20-50	50-80	80-100
2,5-3	3-3,5	3,5-4	4-4,5	100-150	150-200	200-250	250-300
4,5-5	5-5,5	5,5-6	6-6,5	200-350	350-400	400-500	500-600
6,5-7	> 7 pct.			600-700	> 700 kr.		

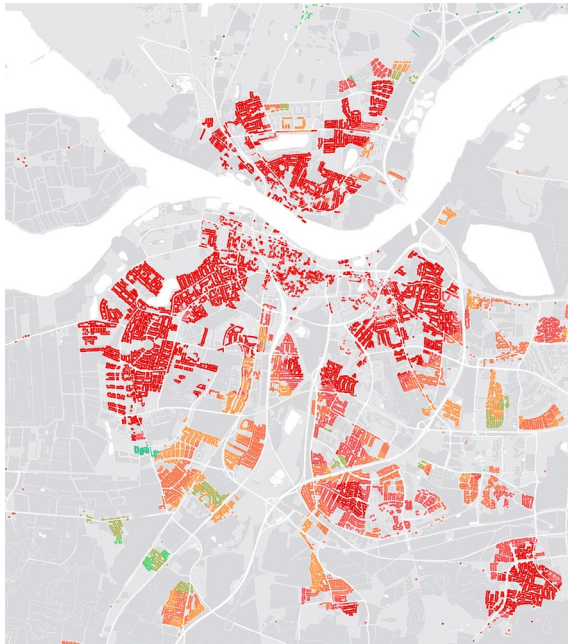
Anm.: Figuren viser den estimerede gennemsnitlige årlige ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

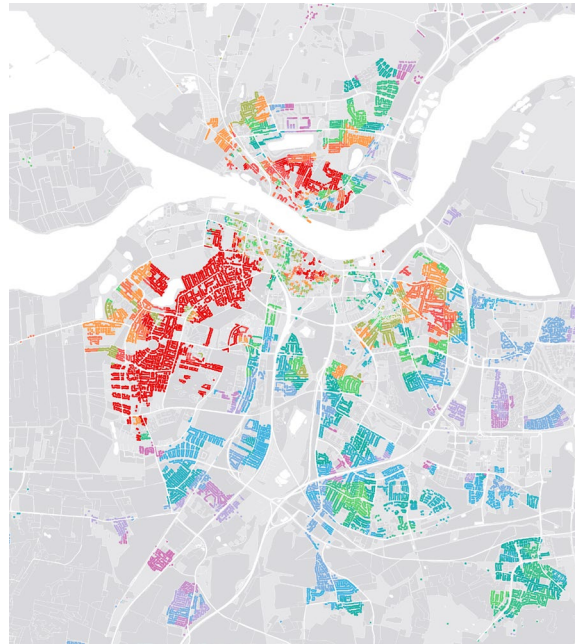
**Figur B.9**

Gennemsnitlig årlig ændring i boligprisen pr. kvadratmeter, Aalborg, 1999-2006 og 2013-2021

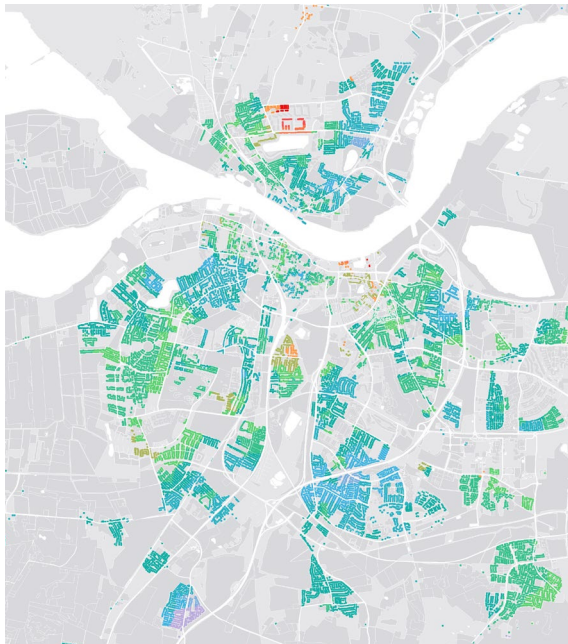
a) Gns. årlig ændring 1999-2006, pct.



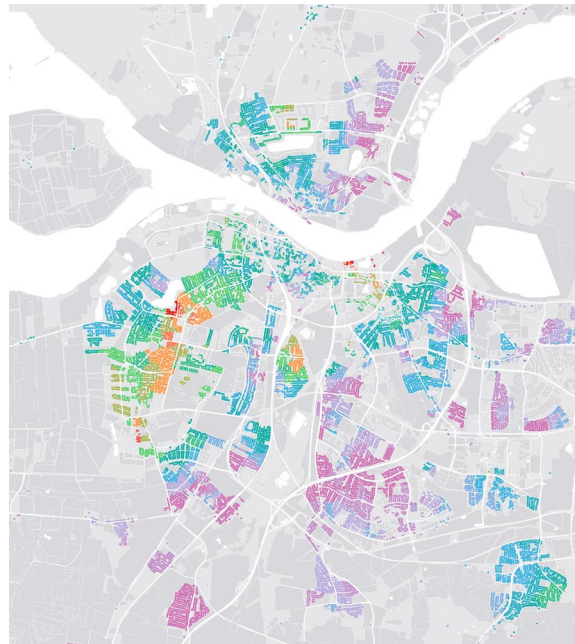
b) Gns. årlig ændring 1999-2006, kr. (2021-niveau)



c) Gns. årlig ændring 2013-2021, pct.



d) Gns. årlig ændring 2013-2021, kr. (2021-niveau)



< 1 pct.	1-1,5	1,5-2	2-2,5	< 400 kr.	400-450	450-500	500-550
2,5-3	3-3,5	3,5-4	4-4,5	550-600	600-650	650-700	700-750
4,5-5	5-5,5	5,5-6	6-6,5	750-800	800-850	850-900	900-950
6,5-7	> 7 pct.			950-1.000	> 1.000 kr.		

Anm.: Figuren viser den estimerede gennemsnitlige årlige ændring i kvadratmeterprisen for en standard ejerbolig beliggende, hvor prikken er placeret.

Kilde: Egne beregninger på baggrund af registerdata fra Danmarks Statistik.

## C. Undersøgelsens datagrundlag

Undersøgelsens datagrundlag er en række administrative registre, som indeholder oplysninger på befolkningen og boligerne i Danmark. Registerne er tilgængelige for DREAM under Danmarks Statistiks forskerserviceordning.

Antallet af boliger i Danmark opgøres på baggrund af Boligopgørelsen, jf. Danmarks Statistik (2023). Dette er en totaltælling af samtlige boliger i Danmark pr. 1. januar i perioden 1992-2022. Opgørelsen pr. 1. januar 2022 omtales typisk som primo 2022 eller ultimo 2021.

Boligopgørelsen baseres på de administrative registre BBR (Bygnings- og Boligregistret) samt CPR (Det Centrale Personregister). BBR er et landsdækkende register med data om samtlige landets bygninger og boliger, mens CPR indeholder oplysning om hver danskers navn, køn, bopælsadresse, civilstand mv.

Ejerregistret gør det muligt at identificere ejere af hver enkelt ejerbolig. Herved bliver det muligt at finde baggrundskarakteristika på den enkelte boligejer. Baggrundskarakteristika omfatter blandt andet alder, identifikation af husstandsmedlemmer, indkomstforhold mv. Salgsoplysninger på solgte ejerboliger findes i Danmarks Statistiks register for ejendomssalg.

### C.1 Ejendomssalg

Danmarks Statistik danner statistikken ejendomssalg, jf. Danmarks Statistik (2023). Registret er baseret på salgsoplysninger fra tinglysningssystemet. Registret begrænses til kun at indeholde salg af ejerboliger. Ved salg af ejerboliger anvendes kun salg i almindelig fri handel og der anvendes samme metode som Danmarks Statistik til at frasortere for ejendomssalg med en atypisk prisudvikling. Frasorteringen omhandler blandt andet boliger med en ekstrem købesum. Ud over kriterierne anvendt af Danmarks Statistik er det i nærværende undersøgelse en forudsætning, at den handlede ejerbolig er privatejet før salget, så ejernes karakteristika kan observeres.

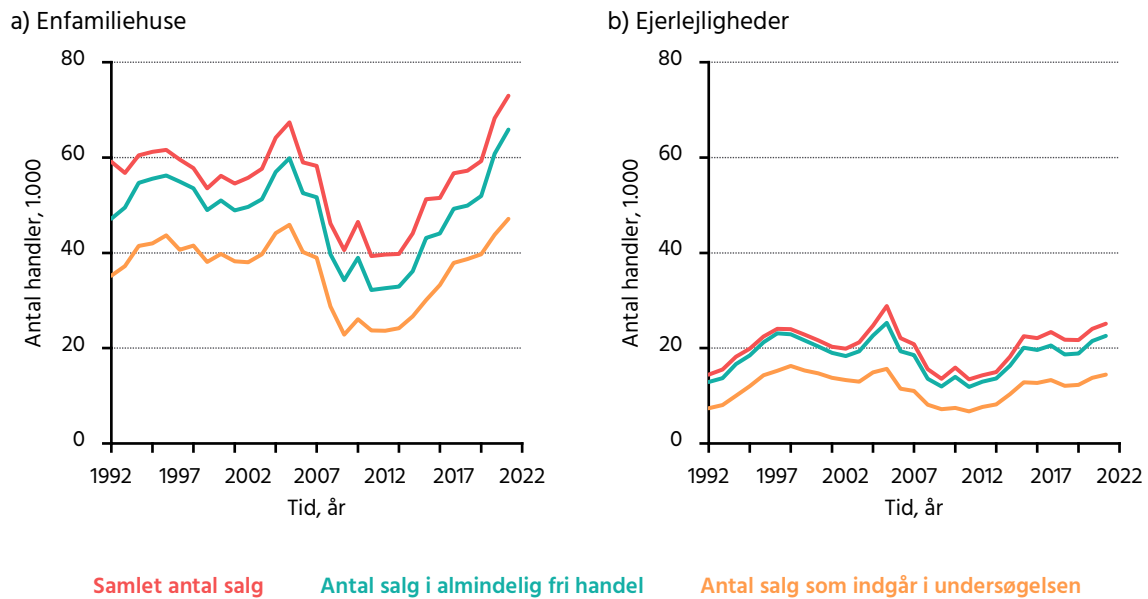
Statistikken omfatter tinglyste bolighandler. Idet tinglysning kan finde sted et stykke tid efter, handlen er gennemført, går der en vis tid, inden statistikken omfatter alle handler og derved kan betragtes som endelig. Dette er årsagen til, at kun bolighandler til og med 2021 betragtes i nærværende undersøgelse.

For perioden 1992-2021 indgår i alt 1,4 mio. handler med ejerboliger i undersøgelsen. Heraf er 1.082.000 enfamiliehuse og 352.000 er ejerlejligheder. Antal bolighandler i det enkelte år svinger med konjunkturerne, jf. Figur C.1.

I undersøgelsen indgår cirka tre fjerdedele af de enfamiliehuse, som handles i fri handel. Andelen er nogenlunde konstant over tid. De handler som frasorteres kan eksempelvis have en ekstrem købesum eller boligen er ikke privatejet før salget. For ejerlejligheder er den tilsvarende andel cirka 65 pct.

Figur C.1

Antal bolighandler



Anm.: Samlet antal salg indeholder salg i almindelig fri handel, familieoverdragelser og øvrige salg herunder tvangsauktioner.  
Kilde: Danmarks Statistik og egne beregninger.

## C.2 Sammenligning med øvrige boligprisindeks

Flere institutioner laver indeks for boligprisudviklingen i Danmark. Udover nærværende undersøgelse omfatter dette blandt andet Danmarks Statistik (ejendomssalgstatistikken), Finans Danmark (boligmarkedsstatistikken) og Boligøkonomisk Videncenter. Desuden vurderer Boligsiden udviklingen i kvadratmeterprisen, dog primært i nyere tid, hvorfor lange tids-serier ikke er tilgængelige.

Generelt gælder det, at de fire metoder vurderer boligprisudviklingen tilnærmelsesvis ens, jf. Figur C.2 og Figur C.3. Der ses en tendens til, at metoden anvendt i nærværende undersøgelse resulterer i en mere glat udvikling med mindre udsving fra kvartal til kvartal, end det er tilfældet ved de øvrige boligprisindeks.

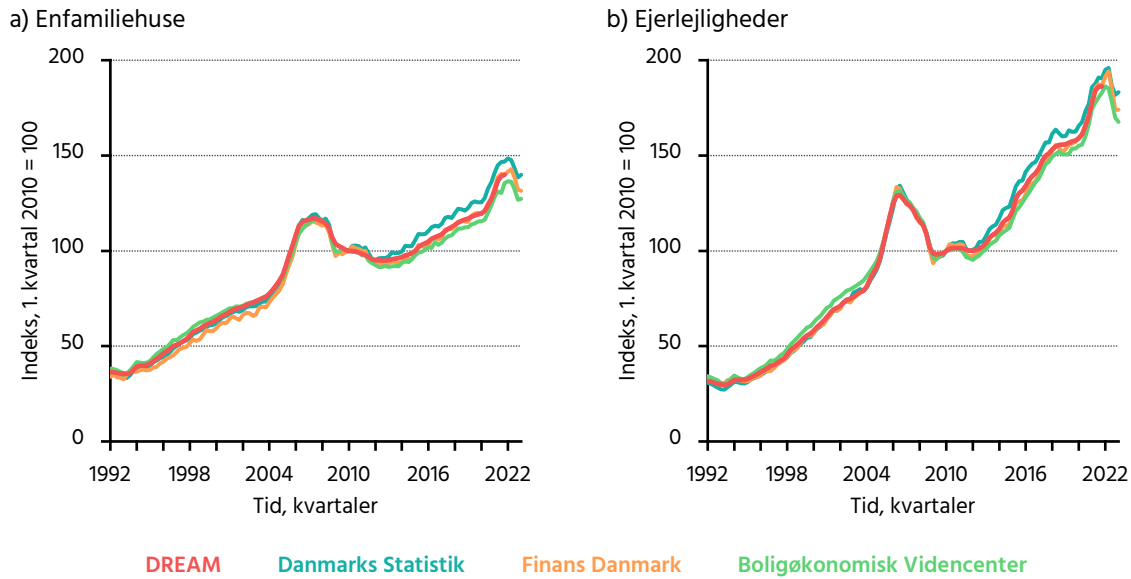
Fælles for de fire førstnævnte institutioner er, at prisindekset videst muligt baseres på bolighandler, hvor handelsprisen fastsættes på markedsvilkår. Der anvendes imidlertid fire forskellige metoder til at vurdere udviklingen i boligprisen. Metoderne skitseres i det følgende.

I nærværende undersøgelse anvendes en metode, som kvartalvis vurderer kvadratmeterprisen for hver enkelt ejerbolig. Herudfra kan gennemsnits- eller medianprisen for et område fastsættes, og prisudviklingen over tid kan betragtes. Metoden til fastsættelse af boligprisen beskrives i kapitel 5.

Danmarks Statistiks boligprisindeks anvender en boligs handelspris og boligens offentlige ejendomsvurdering. Forholdet mellem disse kaldes afstandsprocenten. Hvis afstandsprocenten i to perioder beregnes på baggrund af samme ejendomsvurdering, kan man bruge forskellen på de to perioders afstandsprocenter til at angive prisudviklingen mellem de to perioder. Metoden forudsætter, at der for alle landets ejendomme foreligger en ensartet offentlig prisvurdering. Metoden kaldes Sales Price Appraisal Ratio forkortet SPAR.

Figur C.2

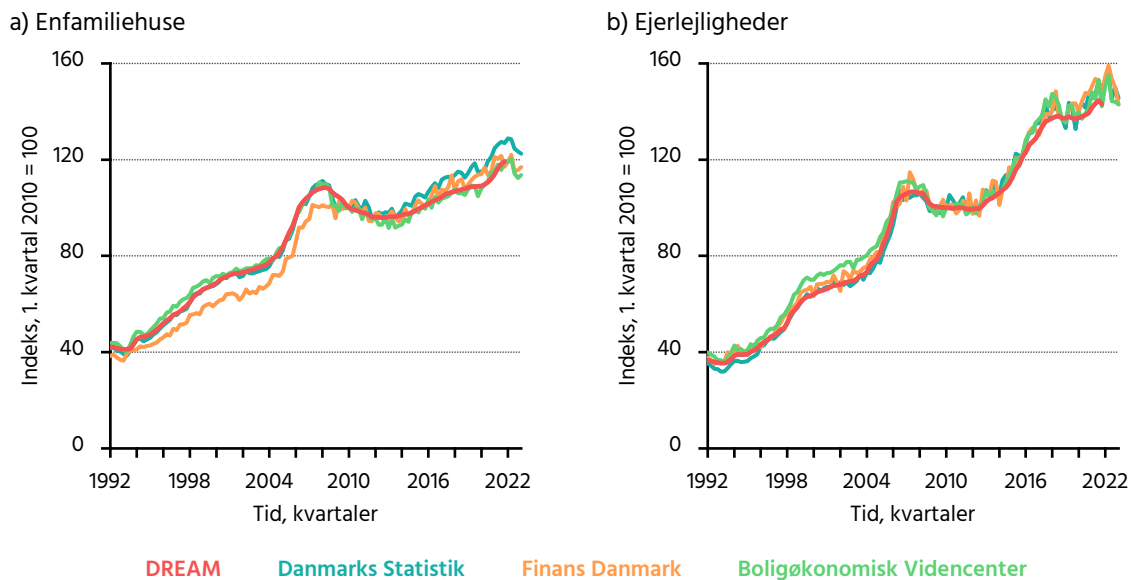
Sammenligning af boligprisindeks, hele landet



Kilde: Egne beregninger, Danmarks Statistik ejendomssalgstatistikken, Finans Danmark boligmarkedsstatistikken og Boligøkonomisk Videncenters boligprisindeks.

Figur C.3

Sammenligning af boligprisindeks, region Nordjylland



Kilde: Egne beregninger, Danmarks Statistik ejendomssalgstatistikken, Finans Danmark boligmarkedsstatistikken og Boligøkonomisk Videncenters boligprisindeks.

Finans Danmark vurderer udviklingen i den gennemsnitlige salgspris pr. boligkvadratmeter i forskellige dele af landet. Der benyttes en vægtet kvadratmeterpris. Dels anvendes boligens vægtede etageareal, hvor for eksempel garage og kælder indgår men vejer mindre end det

egentlige boligareal. Dels vægtes salgspriserne så der korrigeres for forskelle mellem handlede boliger og den samlede boligbestand. En lignende metode anvendes ved dannelse af Boligsidens Markedsindeks.

Boligøkonomisk Videncenter anvender gentagne salgs metode til deres boligprisindeks (eng: repeat sales method). Metoden måler prisudviklingen ved at betragte boliger, som er handlet flere gange. Herved måles prisudviklingen for den samme bolig over tid. Det forsøges at frasortere boliger, som er blevet renoveret, ligesom boliger, der kun er solgt én gang siden 1992, ikke indgår i beregningen.

### Ejerboligpris for kommuner og postnumre

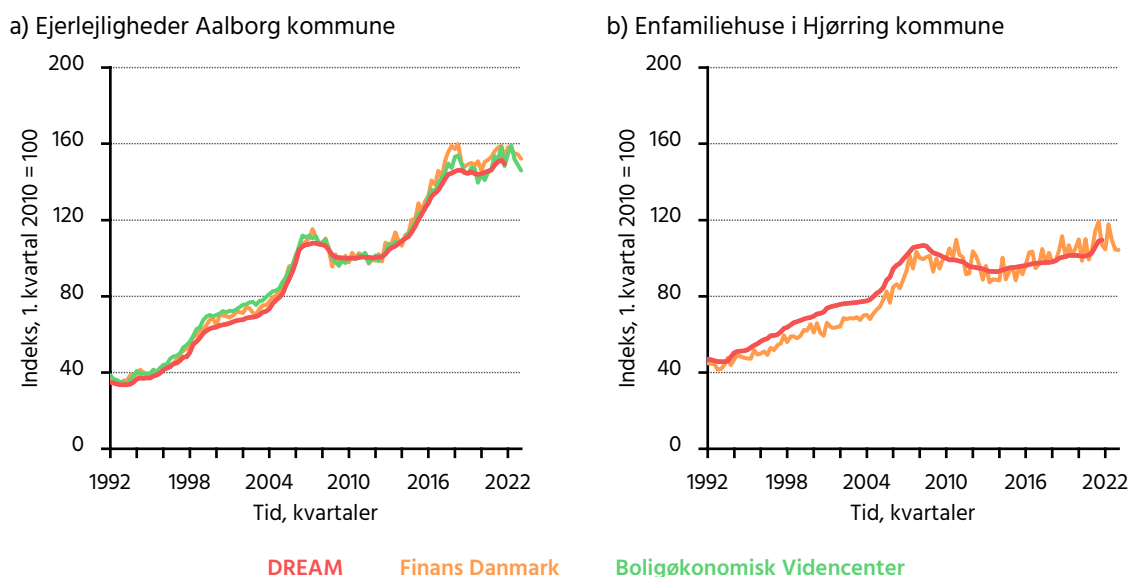
Danmarks Statistik offentliggør boligprisindeks for hele landet, regioner og landsdele. Det samme gør de øvrige aktører.

I tillæg hertil har Boligøkonomisk Videncenter boligprisindeks for kommunerne med landets fire største byer. Både Finans Danmark og Boligsiden fastlægger kvadratmeterpriser på kommuneniveau, hvis der i et givent kvartal er handlet mindst fem boliger i kommunen. Boligsiden har herudover gennemsnitlige boligpriser for enfamiliehuse i cirka 150 postnumre og ejerlejligheder i cirka 60 postnumre. Finans Danmark offentliggør gennemsnitlige kvadratmeterpriser for samtlige postnumre, forudsat at der er sket mindst fem handler i kvartalet.

En fordel ved nærværende metode er, at den for hvert kvartal giver et kvalificeret bud på kvadratmeterprisen for samtlige ejerboliger i landet. Dette gør det for eksempel muligt at vurdere prisudviklingen inden for samtlige kommuner og postnumre uanset, om der er handlet boliger eller ej.

Figur C.4 og Figur C.5 viser eksempler på udviklingen i ejerboligprisen på kommunalt plan og for to postnumre med relativt mange ejerboliger. I postnummer 9400 Nørresundby ligger godt 5.000 ejerboliger, hvoraf cirka 2.000 er ejerlejligheder. I postnummer 9850 Hirtshals er i omegnen af 3.500 ejerboliger. Heraf er 93 pct. enfamiliehuse, og de resterende er lejligheder.

**Figur C.4**  
Sammenligning af boligprisindeks, udvalgte kommuner



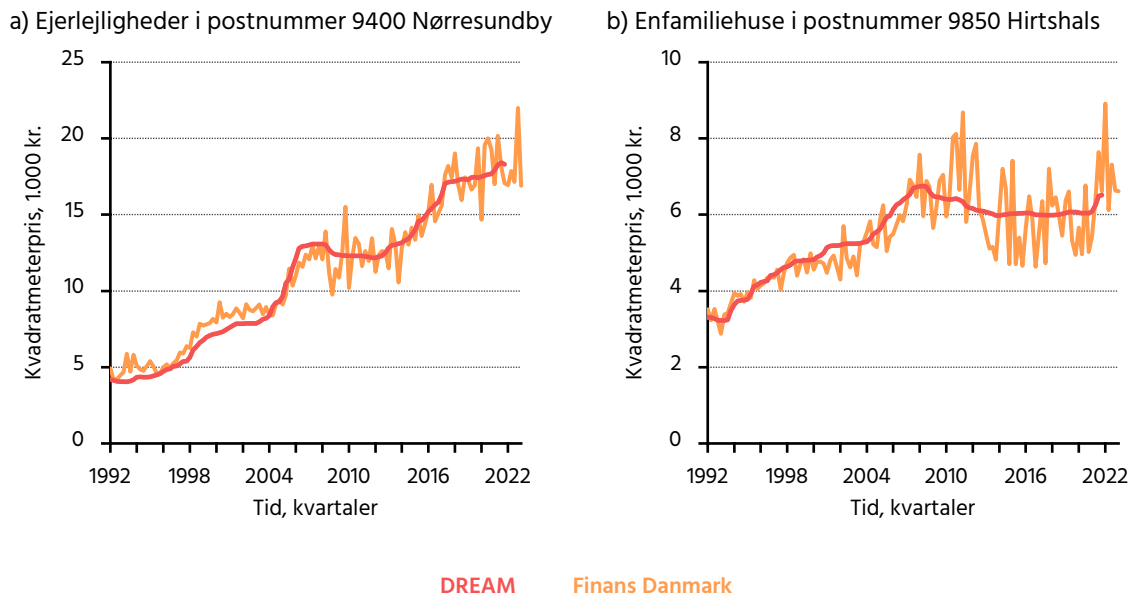
Anm.: Danmarks Statistik offentliggør ikke boligprisudvikling på kommuneniveau, hvorfor denne institution ikke indgår i figuren. Boligøkonomisk Videncenter danner prisudvikling for de fire største byer, hvorfor denne kun vises i figur a).

Kilde: Egne beregninger, Finans Danmark boligmarkedsstatistikken og Boligøkonomisk Videncenters boligprisindeks.



Figur C.5

Sammenligning af boligprisindeks, udvalgte postnumre



Anm.: Danmarks Statistik og Boligøkonomisk Videncenter offentliggør ikke boligprisudvikling på postnumre, hvorfor disse institutioner ikke indgår i figuren.

Kilde: Egne beregninger og Finans Danmark boligmarkedsstatistikken.

### C.3 Familieækvivalering af indkomst

I undersøgelsen er det anvendte indkomstbegreb den familieækvivalerede disponible indkomst. Først opgøres en persons disponible indkomst uden boligrelaterede udgifter. Efterfølgende familieækvivaleres dette indkomstbegreb.

En persons disponible indkomst defineres som summen af arbejds-, kapital- og overførselsindkomster samt udbetaling af private og arbejdsgiveradministrerede pensioner fratrukket personlige skatter.

Arbejdsindkomst består af lønindkomst og overskud af egen virksomhed, hvor lønindkomst er skattepligtig løn inkl. frynsegoder, skattefri løn, jubilæums- og fratrædelsesgodtgørelser samt værdi af aktieoptioner, løn under sygdom og honorar for bestyrelsesarbejde.

Kapitalindkomst er realiserede kapitalindkomster, mens overførselsindkomster er offentlige indkomstoverførsler som f.eks. kontanthjælp, offentlige pensioner, dagpenge og børnefamilieydelse.

Den disponible indkomst familieækvivaleres for at gøre indkomsterne sammenlignelige på tværs af familier med forskellig størrelse. Derfor korrigeres den samlede disponible indkomst med den såkaldte ækvivaleringsfaktor. Den anvendte ækvivaleringsfaktor er givet som i De Økonomiske Råd (2016, s. 212).

Den familieækvivalerede disponible indkomst findes ved at dividere familiens samlede indkomst med ækvivaleringsfaktoren beregnet som  $(\text{antallet af voksne} + 0,6 \cdot \text{antallet af børn})^{0,8}$ .

Hver enkelt person i husholdningen får tildelt en indkomst, der er lig husstandens samlede indkomst divideret med ækvivaleringsfaktoren. Det betyder, at fordelingen af den ækvivalerede indkomst er mere lige end fordelingen af den personlige indkomst. For en enlig er ækvivaleringsfaktoren 1, mens den for en familie med to voksne er 1,74, og for en familie med to voksne og to børn er den 2,54.

## C.4 Bystørrelseskode, landsdele og kommune-gruppe

I rapporten anvendes fem forskellige geografiske opdelinger af landet:

- Den administrative enhed indeholdende fem regioner,
- En statistisk inddeling i 11 landsdele, som er en underopdeling af de fem regioner,
- Den administrative enhed med 98 kommuner,
- En typebestemmelse af hver kommune som økommune, landkommune, oplandskommune, provinskommune, storbykommune eller hovedstadskommune ud fra antallet af indbyggere i kommunens største by og tilgængelighed til jobs,
- En opdeling efter størrelse af det byområde, hvor en given bolig er beliggende.

I dette afsnit beskriver de tre inddelinger, som ikke er administrative enheder. Tabel C.1 viser sammenhængen mellem regioner, landsdele og kommuner.

**Tabel C.1**  
Sammenhæng mellem regioner, landsdele og kommuner

Region	Landsdel	Kommuner
Hovedstaden	København By	Københavns, Frederiksberg, Dragør og Tårnby kommuner
	Københavns omegn	Albertslund, Ballerup, Brøndby, Gentofte, Gladsaxe, Glostrup, Herlev, Hvidovre, Høje-Taastrup, Ishøj, Lyngby-Taarbæk, Rødovre og Vallensbæk kommuner
	Nordsjælland	Allerød, Egedal, Fredensborg, Frederikssund, Furesø, Gribskov, Halsnæs, Helsingør, Hillerød, Hørsholm og Rudersdal kommuner
	Bornholm	Bornholms Kommune og Ertholmene
Sjælland	Østsjælland	Greve, Køge, Lejre, Roskilde og Solrød kommuner
	Vest- og Sydsjælland	Faxe, Guldborgsund, Holbæk, Kalundborg, Lolland, Næstved, Odsherred, Ringsted, Slagelse, Sorø, Stevns og Vordingborg
Syddanmark	Fyn	Assens, Faaborg-Midtfyn, Kerteminde, Langeland, Middelfart, Nordfyn, Nyborg, Odense, Svendborg og Ærø kommuner
	Sydjylland	Billund, Esbjerg, Fanø, Fredericia, Haderslev, Kolding, Sønderborg, Tønder, Varde, Vejen, Vejle og Aabenraa kommuner
Midtjylland	Vestjylland	Herning, Holstebro, Ikast-Brande, Lemvig, Ringkøbing-Skjern, Skive, Struer og Viborg kommuner
	Østjylland	Favrskov, Hedensted, Horsens, Norddjurs, Odder, Randers, Samsø, Silkeborg, Skanderborg, Syddjurs og Aarhus kommuner
Nordjylland	Nordjylland	Brønderslev, Frederikshavn, Hjørring, Jammerbugt, Læsø, Mariagerfjord, Morsø, Rebild, Thisted, Vesthimmerland og Aalborg

Anm.: Regioner og kommuner er administrative inddelinger, mens landsdele er dannet til statistiske formål.  
Kilde: Danmarks Statistik.

Danmarks Statistik foretager en inddeling af landets kommuner, hvor kommunerne grupperes efter den største by i kommunen og tilgængelighed til job.

Kriterierne er valgt for at kunne reflektere to forhold. For det første hvor vidt en given kommune har en by af en vis størrelse (by-land dimension). For det andet om kommunen ligger 'centralt' eller 'yderligt' i forhold til tilgængelighed til arbejdspladser. Jobtilgængeligheden udtrykker hvor mange jobs, der kan nås fra en given kommune, idet antallet af jobs vægtes negativt med afstanden til arbejdspladsen.

Kommunegrupperingen foretages successivt efter følgende kriterier:

1. Kommuner med jobtilgængelighed på mindst 200.000 karakteriseres som hovedstadskommuner.
2. Kommuner med mindst 100.000 indbyggere i største by kaldes storbykommuner.
3. Kommuner med mindst 30.000 indbyggere i den største by kaldes provinsbykommuner.
4. Hvis tilgængeligheden til arbejdspladser er mindst 40.000 karakteriseres som oplandskommuner.
5. Kommuner som ikke placeres i kategoriseres under punkt 1. til 4. kaldes land- eller økommuner. Sidstnævnte gruppe omfatter Bornholm, Ærø, Fanø og Læsø kommuner.

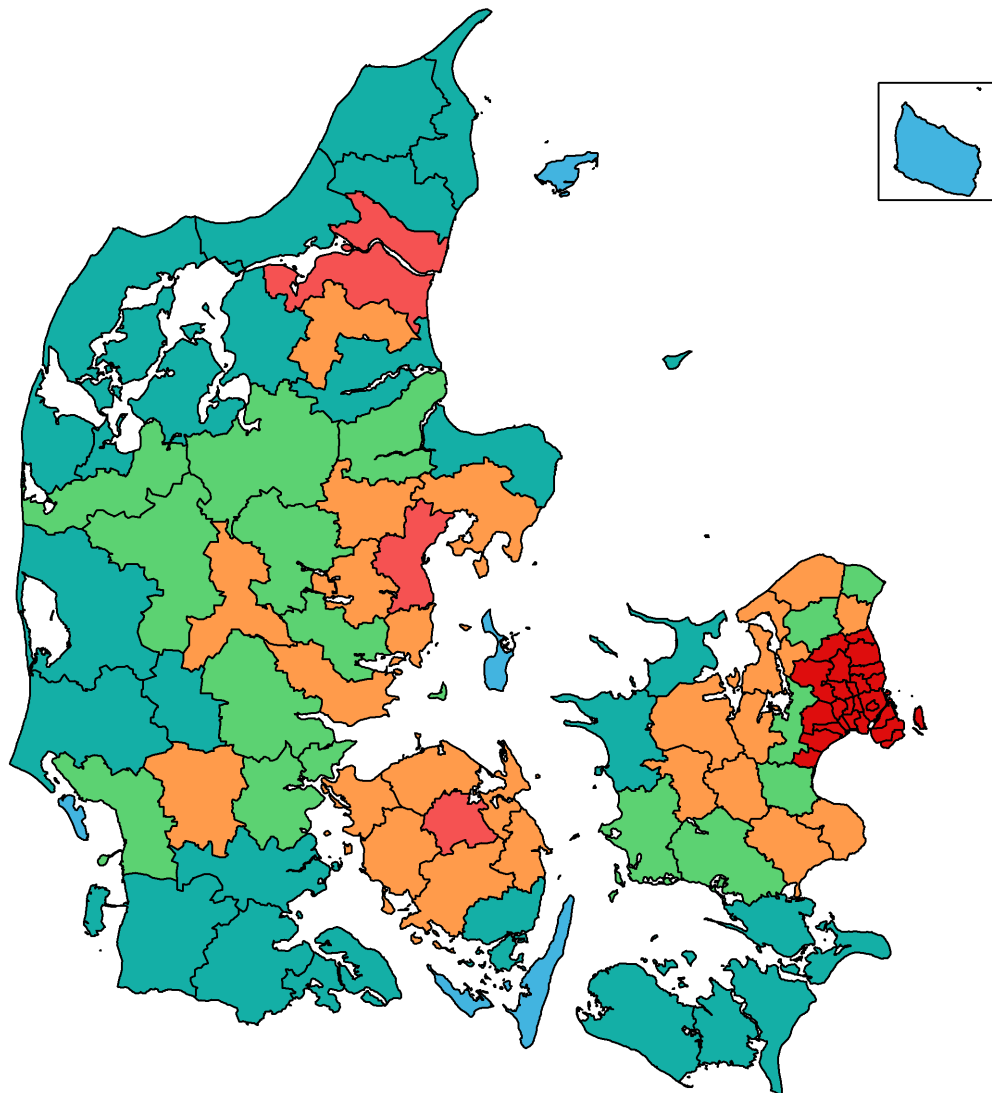
Kommunegrupperingen er vist i Tabel C.2 og Figur C.5. Opgørelsesmetoden er dokumenteret i Danmarks Statistik (2018).

**Tabel C.2**  
**Definition af kommunegrupper**

Kommunegruppe	Kommuner
Hovedstadskommuner	København, Frederiksberg, Ballerup, Brøndby, Dragør, Gentofte, Gladsaxe, Glostrup, Herlev, Albertslund, Hvidovre, Høje-Taastrup, Lyngby-Taarbæk, Rødovre, Ishøj, Tårnby, Valensbæk, Furesø, Allerød, Hørsholm, Rudersdal, Egedal, Greve og Solrød
Storbykommuner	Odense, Aarhus og Aalborg
Provinsbykommuner	Helsingør, Hillerød, Køge, Roskilde, Slagelse, Næstved, Esbjerg, Fredericia, Horsens, Kolding, Vejle, Herning, Holstebro, Randers, Silkeborg og Viborg
Oplandskommuner	Fredensborg, Frederikssund, Halsnæs, Gribskov, Holbæk, Faxe, Ringsted, Stevns, Sorø, Lejre, Middelfart, Assens, Faaborg-Midtfyn, Kerteminde, Nyborg, Nordfyns, Vejen, Syddjurs, Favrskov, Odder, Skanderborg, Ikast-Brande, Hedensted og Rebild
Landkommuner	Odsherred, Kalundborg, Lolland, Guldborgsund, Vordingborg, Svendborg, Haderslev, Billund, Sønderborg, Tønder, Varde, Aabenraa, Lemvig, Struer, Norddjurs, Ringkøbing-Skjern, Morsø, Skive, Thisted, Brønderslev, Frederikshavn, Vesthimmerlands, Mariagerfjord, Jammerbugt og Hjørring
Økommuner	Bornholm, Langeland, Ærø, Fanø, Samsø og Læsø

Anm.: Gruppen indeholdende økommuner indgår ikke i Danmarks Statistiks gruppering. Gruppen er dannet af DREAM.  
Kilde: Danmarks Statistik og egen opgørelse.

**Figur C.6**  
Visualisering af kommune-grupper



**Hovedstadskommune** **Storbykommune** **Provinsbykommune** **Oplandskommune** **Landkommune** **Økommune**

Anm.: Gruppen indeholdende økommuner indgår ikke i Danmarks Statistiks gruppering. Gruppen er dannet af DREAM.

Kilde: Danmarks Statistik og egen opgørelse.

En by defineres som en naturligt sammenhængende bebyggelse med mindst 200 indbyggere.

At bebyggelsen er naturligt sammenhængende vil sige, at afstanden mellem husene normalt ikke overstiger 200 meter medmindre afbrydelsen skyldes større gennemgående veje (uden direkte adgangsveje mellem bebyggelserne), kirkegårde, sportspladser, parkerings- og parkanlæg, jernbane- og lagerarealer, jorde under udstykning eller lignende.

Bebyggelser med under 200 indbyggere betragtes som landdistrikter. Spredt bebyggelse langs en landevej henregnes ikke som hørende til en by, heller ikke hvis afstanden imellem husene er mindre end 200 meter.

For at to byer kan klassificeres som vokset sammen forudsætter det både, at afstanden mellem dem er under 200 meter samt, at de ud fra et topografisk skøn hører naturligt sammen. Tabel C.3 viser den anvendte gruppering af bystørrelser samt kategoriserede bynavne på byer med mindst 10.000 indbyggere.

**Tabel C.3**

**Byer efter bystørrelseskode**

Bystørrelseskode	Afgrænsning af kategorien
Hovedstadsområdet	København, Frederiksberg, Albertslund, Brøndby, Gentofte, Gladsaxe, Glostrup, Herlev, Hvidovre, Lyngby-Taarbæk, Rødovre, Tårnby og Vallensbæk kommuner, dele af Ballerup, Rudersdal og Furesø, samt Ishøj By og Greve Strand By.
<b>Byområder i øvrigt med</b>	
Mindst 100.000 indbyggere	Aarhus, Odense, Aalborg
50.000-99.999 indbyggere	Esbjerg, Randers, Kolding, Horsens, Vejle, Roskilde, Herning
40.000-49.999 indbyggere	Silkeborg, Helsingør, Næstved, Viborg, Fredericia samt dele af Rudersdal, Hørsholm og Fredensborg kommuner (sammenhængende byområder langs kysten)
30.000-39.999 indbyggere	Køge, Holstebro, Høje-Taastrup, Slagelse, Hillerød
20.000-29.999 indbyggere	Holbæk, Sønderborg, Svendborg, Hjørring, Nørresundby, Ringsted, Frederikshavn, Ølstykke-Stenløse, Haderslev, Birkerød, Skive, Farum samt dele af Ballerup/Egedal kommuner (området omkring Smørum/Måløv)
10.000-19.999 indbyggere	Skanderborg, Nyborg, Solrød Strand, Nykøbing F, Lillerød, Frederikssund, Aabenraa, Kalundborg, Middelfart, Ikast, Korsør, Grenaa, Varde, Rønne, Thisted, Hedehusene, Frederiksværk, Nakskov, Brønderslev, Odder, Dragør, Hedensted, Haslev, Hobro, Vordingborg, Lystrup, Jyllinge
5.000-9.999 indbyggere	-
2.000-4.999 indbyggere	-
1.000-1.999 indbyggere	-
500-999 indbyggere	-
200-499 indbyggere	-
<b>Landdistrikter</b>	
Færre end 200 indbyggere	-

Anm.: Bystørrelser er opgjort ultimo 2020.  
Kilde: Danmarks Statistiks byopgørelse.

# DREAM

Danish Research Institute for  
Economic Analysis and Modelling

DREAM

Landgreven 4

1301 København K

Tlf.: 25 26 27 60

[info@dreamgruppen.dk](mailto:info@dreamgruppen.dk)

[www.dreamgruppen.dk](http://www.dreamgruppen.dk)

