

Datagrundlag for Landsbyggefondens Tvillingeværktøj

Data i Tvillingeværktøjet

Der indgår en række forskellige data i Tvillingeværktøjet. Data opdateres årligt, når der foreligger nye data. Dette gælder både regnskabsdata, effektivitetsdata, stamdata og socioøkonomiske data (se Tabel 4).

Tabel 1: Data for følgende baggrundsvARIABLE indgår i udsøgningen af tvillinger på boligorganisationsniveau:

BaggrundsvARIABLE	Definition af baggrundsvARIABLE	Datakilde
Antal lejemålsenheder	Lejemålsenheder er summen af antal boliglejemål (1 bolig = 1 lejemålsenhed.), erhvervslejemål og institutioner (1 lejemålsenhed pr. påbegyndt 60 m²) og garager/carporte (tæller 1/5 lejemålsenhed). Summeret på boligorganisationsniveau.	Landsbyggefondens Stamdatabase + Regnskabsdatabase
Gennemsnitlig afdelingsstørrelse	Antal lejemålsenheder/antal afdelinger	Landsbyggefondens Stamdatabase + Regnskabsdatabase
Andel indvandrere/- efterkommere fra ikke-vestlige lande	Det samlede antal indvandrere og efterkommere fra ikke-vestlige lande i afdelingen set i forhold til det samlede antal beboere. Summeret på boligorganisationsniveau.	Danmarks Statistik
Andel udenfor arbejdsmarkedet	Antallet af fuldtidspersoner mellem 16-64 år udenfor arbejdsmarkedet set ift. antallet af beboere i afdelingen mellem 16 og 64 år, begge aldre inkl. Summeret på boligorganisationsniveau. Beboere udenfor arbejdsmarkedet er defineret som personer (omregnet til fuldtidspersoner) som tilhører én af nedenstående 5 grupper: Gruppe A: Samtlige registrerede nettoledige (inkl. dagpengemodtagere, arbejdsmarkedsydelse og kontanthjælps-/uddannelseshjælps-/ og integrationsydelsesmodtagere, der visiteres som jobparate eller åbenlyst uddannelsesparate). Gruppe B: Samtlige aktiverede på offentlig forsørgelse, dog ikke aktiverede på sygedagpenge. Herudover samtlige modtagere af revalideringsydelse og ledighedsydelse. Gruppe C: Passive kontanthjælpsmodtagere og uddannelseshjælpsmodtagere (uden ledighed eller aktivering) samt integrationsydelsesmodtagere (uden ledighed eller aktivering). Herudover samtlige personer i ressourceforløb og jobafklaring (uden aktivering). Gruppe D: Personer på sygedagpenge fra ledighed herunder aktiverede på sygedagpenge fra ledighed. Gruppe E: Personer på førtidspension (herunder også personer i skånejob).	Danmarks Statistik

Indkomst ift. gennemsnitlig landsindkomst	Kun fuldt skattepligtige personer på 15 år og derover er medtaget i beregningen. Anvendt indkomst er personindkomst i alt ekskl. beregnet lejeværdi af egen bolig og før fradrag af renteudgifter. Hvis afdelingen/organisationens gennemsnitlige indkomst pr. person er større end landsgennemsnittet er værdien af andelen større end 1. Er dette tilfældet er andelen sat til 1. Er afdelingens/organisations gennemsnitlige indkomst pr. person under 25 % af landsgennemsnittet sættes den til 25 %.	Danmarks Statistik
Andel personer uden uddannelse/uoplyst uddannelse	Det samlede antal personer på 15 år og derover med uoplyst uddannelse eller grundskole som højeste fuldførte uddannelse set ift. det samlede antal beboere på 15 år og derover. Personer under uddannelse er ikke medtaget. Summeret på boligorganisationsniveau.	Danmarks Statistik

Tabel 2: Data for følgende baggrundsvARIABLE indgår i udsøgningen af tvillinger på afdelingsniveau:

BaggrundsvARIABLE	Definition af baggrundsvARIABLE	Datakilde
Antal lejemålsenheder	Lejemålsenheder er summen af antal boliglejemål (1 bolig = 1 lejemålsenhed.), erhvervslejemål og institutioner (1 lejemålsenhed pr. påbegyndt 60 m ²) og garager/carporte (tæller 1/5 lejemålsenhed).	Landsbyggefondens Stamdatabase + Regnskabsdatabase
Ibrugtagelsesår	Afdelingers ibrugtagelsesår opgøres efter ibrugtagelsesåret for det første ibrugtagne byggeafsnit. Afdelinger uden tilknyttet ibrugtagelsesår indgår ikke.	Landsbyggefondens Stamdatabase
Andel indvandrere/efterkommere fra ikke-vestlige lande	Det samlede antal indvandrere og efterkommere fra ikke-vestlige lande i afdelingen set i forhold til det samlede antal beboere i afdelingen.	Danmarks Statistik
Andel udenfor arbejdsmarkedet	Antallet af fuldtidspersoner mellem 16-64 år udenfor arbejdsmarkedet i afdelingen set ift. antallet af beboere i afdelingen mellem 16 og 64 år, begge aldre inkl. Beboere udenfor arbejdsmarkedet er defineret som personer (omregnet til fuldtidspersoner) som tilhører én af nedenstående 5 grupper: Gruppe A: Samtlige registrerede nettoledige (inkl. dagpengemodtagere, arbejdsmarkedsydelse og kontanthjælps-/uddannelseshjælps- og integrationsydelsesmodtagere, der visiteres som jobparate eller åbenlyst uddannelsesparate).	Danmarks Statistik

	Gruppe B: Samtlige aktiverede på offentlig forsørgelse, dog ikke aktiverede på sygedagpenge. Herudover samtlige modtagere af revalideringsydelse og ledighedsydelse. Gruppe C: Passive kontanthjælpsmodtagere og uddannelseshjælpsmodtagere (uden ledighed eller aktivering) samt integrationsydelsesmodtagere (uden ledighed eller aktivering). Herudover samtlige personer i ressourceforløb og jobafklaring (uden aktivering). Gruppe D: Personer på sygedagpenge fra ledighed herunder aktiverede på sygedagpenge fra ledighed. Gruppe E: Personer på førtidspension (herunder også personer i skånejob).	
Indkomst ift. gennemsnitlig landsindkomst	Kun fuldt skattepligtige personer på 15 år og derover er medtaget i beregningen. Anvendt indkomst er personindkomst i alt ekskl. beregnet lejeværdi af egen bolig og før fradrag af renteudgifter. Hvis afdelingen/organisationens gennemsnitlige indkomst pr. person er større end landsgennemsnittet er værdien af andelen større end 1. Er dette tilfældet er andelen sat til 1. Er afdelingens/organisationens gennemsnitlige indkomst pr. person under 25 % af landsgennemsnittet sættes den til 25 %.	Danmarks Statistik
Andel personer uden uddannelse/uoplyst uddannelse	Det samlede antal personer på 15 år og derover med uoplyst uddannelse eller grundskole som højeste fuldførte uddannelse set ift. det samlede antal beboere på 15 år og derover. Personer under uddannelse er ikke medtaget. Summeret på boligorganisationsniveau.	Danmarks Statistik

Der vises yderligere en række stamdataoplysninger om boligorganisationerne og afdelingerne, som dog ikke indgår som baggrundsvariable i udsøgningen af tvillinger.

Tabel 3: Yderligere stamdata om boligorganisationer og afdelinger:

Stamdata	Definition	Datakilde
Antal afdelinger i boligorganisationen	Antal boligafdelinger i boligorganisationen	Landsbyggefondens Stamdatabase
Byggeriart	Angiver om en afdeling består af etage eller tæt/lavt byggeri. I afdelinger med begge byggeriarter er det den art med mest boliger, der bestemmer byggeriarten.	Landsbyggefondens Stamdatabase
Antal m ² i afdelingen	Antal m ² i afdelingen	Landsbyggefondens Stamdatabase
Udearealer i afdelingerne	Udearealet er defineret som forskellen mellem det matrikulære areal og det bebyggede areal, for alle matrikler og ejendomme med boliger, som er tilknyttet den enkelte almene boligafdeling. Ejendomme uden boliger fx garager, skure, vaskefaciliteter, fælleslokaler mv. indgår i udearealet. Tilsvarende indgår veje, parkeringspladser o.lign. ligeledes i udearealet i det omfang de ligger på afdelingens matrikler. Ikke matrikuleret areal, som ligger mellem flere af en afdelings matrikler, fx offentlig veje, indgår ikke i afdelingens udearealer. På matrikler med flere ejendomme tilhørende forskellige boligafdelinger, private ejerforeninger eller andre ikke-almene boligformer, er matriklens samlede areal forholdsmæssigt fordelt mellem afdelingerne på baggrund af oplysningerne om antallet af boliger i BBR. Tilsvarende er det bebyggede areal fra ejendomme, som deles af to almene afdelinger eller en almen afdeling og organisationer med andre boligformer, forholdsmæssigt fordelt mellem afdelingerne. Til dette formål er der anvendt en af to vægte, primært ejendommens bebyggede areal og sekundært antallet af boliger i afdelingen / den ikke almene boligform, hvis begge ligger i samme ejendom.	Landsbyggefondens huslejeregister og Landsbyggefondens Stamdatabase koblet med oplysninger fra hhv. BBR og ESR (Det Fælles Kommunale Ejendomsstamregister).

Opdatering af data

Data opdateres årligt, når der foreligger nye data, hvilket forventes at ske ca. i august/september måned hvert år. Dette gælder både regnskabsdata og effektivitetsdata. Stamdata og socioøkonomiske data opdateres i oktober/november.

Tabel 4:

	Hvornår er nuværende data fra?	Hvornår opdateres data næste gang (ca.)?
Regnskabsdata	Regnskabsåret 2023 Opdateres når minimum 80 % af regnskaberne for 2023 er indberettet. ¹	Opdateres til regnskabsåret 2024 ca. august 2025.
Stamdata (lejemålsenheder, afdelingsstørrelse mv.)	Regnskabsperiodens afslutningstidspunkt	Sammen med regnskabsdata
Udearealer	2023	2025
Effektivitetsdata	Effektivitetstal opgjort august 2024, baseret på regnskabsår 2023.	Opdateres med effektivitetstal opgjort august 2025, baseret på regnskabsår 2024, i august/september 2025.
Antal beboere	Data er opgjort pr. 1. januar 2024, baseret på adressegrundlag pr. 1. januar 2024.	Opdateres til data for 1. januar 2025 i oktober/november 2025.
Andel indvandrere/-efterkommere fra ikke-vestlige lande	Data er opgjort pr. 1. januar 2024, baseret på adressegrundlag pr. 1. januar 2024.	Opdateres til data for 1. januar 2025 i oktober/november 2025.
Andel udenfor arbejdsmarkedet	For 2023, baseret på adressegrundlag pr. 1. januar 2024.	Opdateres til data for 2024 i oktober/november 2025.
Indkomst ift. gennemsnitlig landsindkomst	Indkomståret 2023, baseret på adressegrundlag pr. 1. januar 2024.	Opdateres til data for indkomståret 2024 i oktober/november 2025.
Andel personer uden uddannelse/uoplyst uddannelse	Data er opgjort pr. 1. oktober 2023, baseret på adressegrundlag pr. 1. januar 2024.	Opdateres til data for 1. oktober 2024 i oktober/november 2025.

¹ 80 % af regnskaberne (opgjort på boliger). De resterende 20 % opdateres i takt med at de indberettes.

Diskretionering

Visning af socioøkonomiske data på afdelingsniveau omfattes af diskretionering, såfremt afdelingen har mindre end 20 personer mellem 16 og 64 år. Tilsvarende diskretioneres data på boligorganisationsniveau, hvis boligorganisationen har mindre end 20 personer mellem 16-64 år.

Regnskabsdata

Regnskabsnøgletal på boligorganisationsniveau opgøres pr. lejemålsenhed, og på afdelingsniveau opgøres de pr. m².

Regnskabsdatagrundlaget er nærmere beskrevet i "Datagrundlag for Landsbyggefondens regnskabsdatabase", hvorfor der henvises til dette. Datagrundlaget kan findes på fondens hjemmeside eller i regnskabsdatabasen.

Effektivitetsdata

Trafik-, Bygge- og Boligstyrelsens Effektivitetsenhed beregner effektivitetstal. Effektivitetstallene indgår i Almenstyringsdialog og opdateres heri 4 gange om året. Udover et effektivitetstal for de enkelte afdelinger vises også et gennemsnit for boligorganisationen.

I Tvillingeværktøjet indgår effektivitetstallene, der opgøres i august/september.

Der henvises til datagrundlag, metodebeskrivelse mv. på effektiviseringsenhedens hjemmeside.

Beregninger:

Afstandsmål

Tvillingeværktøjet udsøger boligorganisationer/afdelinger der ligner en konkret boligorganisation/afdeling mest muligt ud fra de ovenfor definerede baggrundsvariable. Det sker ud fra en beregning af en "afstand/forskel" mellem den konkrete boligorganisation/afdeling og tvillingerne på de enkelte baggrundsvariable, og vægtet som vist i Tabel 4 og Tabel 5. For en nærmere eksempelbeskrivelse af beregningen af afstandene, se bilag 1.

Tabel 5: Beregning af afstandsmål på boligorganisationsniveau:

Variable	Vægtning
Stamdata/størrelse	De to variable vægter hver 50 %
Socioøkonomiske	De fire variable vægter hver 25 %
Kombineret	Stamdata/størrelse og socioøkonomiske vægter hver 50 %

Tabel 6: Beregning af afstandsmål på afdelingsniveau:

Variable	Vægtning
Stamdata/størrelse	De to variable vægter hver 50 %
Socioøkonomiske	De fire variable vægter hver 25 %
Kombineret	Stamdata/størrelse og socioøkonomiske vægter hver 50 %

Bilag 1: Mere om afstandsmål i tvillingeværktøjet

Afstandsberegningen tager udgangspunkt i følgende afstandsformel (Pythagoras), hvor data ikke er normeret. Nedenstående eksempel er for beregningen af afstanden på de socioøkonomiske variable. Tilsvarende formler anvendes for stamdata/størrelse:

$$\text{Afstand} = \sqrt{w_1 * ((a_1 - a)^2) + w_2 * ((b_1 - b)^2) + w_3 * ((c_1 - c)^2) + w_4 * ((d_1 - d)^2)}$$

a, b, c og d er de 4 socioøkonomiske variable. w1, w2, w3 og w4 angiver vægtingen af de 4 socioøkonomiske variable. Da alle fire variable vægtes lige, er vægtene w1=w2=w3=w4=1. Dermed fås følgende afstandsformel:

$$\text{Afstand} = \sqrt{((a_1 - a)^2) + ((b_1 - b)^2) + ((c_1 - c)^2) + ((d_1 - d)^2)}$$

Hvor	a1	andel IE i afdeling 1.
	a	andel IE i afdeling n.
	b1	andel indkomst i forhold til landsindkomst for afdeling 1.
	b	andel indkomst i forhold til landsindkomst for afdeling n.
	c1	andel personer uden uddannelse/uoplyst uddannelse i afdeling 1.
	c	andel personer uden uddannelse/uoplyst uddannelse i afdeling n.
	d1	andel personer uden tilknytning til arbejdsmarkedet i afdeling 1.
	d	andel personer uden tilknytning til arbejdsmarkedet i afdeling n.

Jo mindre afstand desto mere 'ligner' to afdelinger hinanden.

Eksempel

Lad os antage, der blot eksisterer fem afdelinger. Vi vil gerne finde hver afdelings tætteste tvilling. De fem afdelinger er kendetegnet ved hver to indikatorvariable.

Tabel 1. Indikatorværdier

	Indikator 1	Indikator 2
Afdeling 1	3	21
Afdeling 2	2	22
Afdeling 3	8	21
Afdeling 4	4	23
Afdeling 5	5	23

Først udregnes middelværdien og standardafvigelsen for hver afdeling. Middelværdierne er henholdsvis 3,4 og 22.

Herefter udregnes standardafvigelsen vha. formlen

$$\sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

Hvor x_i angiver indikatorværdien for afdeling i (1-5), $\bar{x}$ angiver middelværdien, mens n angiver det totale antal afdelinger (5).

For indikator 2 er beregningen således

$$\begin{aligned} & \sqrt{\frac{(21-22)^2 + (22-22)^2 + (21-22)^2 + (23-22)^2 + (23-22)^2}{4}} \\ &= \sqrt{\frac{1+0+1+1+1}{4}} \\ &= \sqrt{\frac{4}{4}} = 1 \end{aligned}$$

Standardafvigelsen for indikatorvariabel 2 er dermed 1. På samme måde kan man benytte formlen til at beregne standardafvigelsen for variabel 1. Her fås en standardafvigelsen på 2,3. Standardafvigelsen kaldes også for spredningen og kan intuitivt opfattes som forskellen på indikatorværdien for en tilfældig observation og middelværdien.

Herefter kan t-værdierne beregnes vha. standardafvigelsen. En t-værdi vil praktisk talt altid ligge mellem -3 og 3.

t-værdien regnes med denne formel

$$\frac{x - \bar{x}}{std(x)}$$

Lad os fx regne t-værdien for indikator 1 for afdeling 1.

$$\frac{3 - 4,4}{2,3} = -0,61$$

I nedenstående tabel er alle indikatorværdierne fra tabel 1 omregnet til t-værdier.

t-værdier		
Afdelinger	Indikator 1	Indikator 2
Afdeling 1	-0,60812114	-1
Afdeling 2	-1,04249338	0
Afdeling 3	1,56374007	-1
Afdeling 4	-0,17374889	1
Afdeling 5	0,26062334	1