METODEN BAG ANALYSERNE

Sådan laver vi en analyse: metoden bag "1.000 danske parcelhuse"

Stikprøven, motoren, prisgrundlaget og hvad tallene kan bære · gælder alle analyser

Dette dokument beskriver, hvordan energianalyse.dk regner en stikprøve af danske bygninger igennem og gør resultatet til tal, der kan citeres. Den første analyse (oktober 2026) regnede 1.000 tilfældigt udvalgte fritliggende enfamiliehuse igennem; metoden er bygget, så de næste analyser (etageejendomme, mark-solceller på fri jord) kan følge samme opskrift. Analysens egne tal står på analysesiden og i den anonymiserede CSV; her står kun metoden.

01 Grundprincippet: én sandhed

Hvert tal, der kan citeres, er beregnet af et værktøj og låst af en test. Ingen tal skrives i hånden. Kæden er:

1. **Stikprøven** trækkes af værktøjet med et fast tilfældighedsfrø og gemmes som en udvalgsfil (adresser, kvoter, reserver, vægte).
2. **Beregningen** sker gennem præcis den motor, ejendomsejeren får gratis på siden: hvert hus sendes ind ad den offentlige beregningsvej (ingen parallel model, ingen særlige indstillinger), og motorens fulde svar gemmes som råsvar.
3. **Udtrækket**: hver række i analysens datafil udledes af husets råsvar gennem én funktion, så en kolonne kun kan betyde én ting.
4. **Rapporten**: tal-blokken i analysens markdown, analysesidens JSON og den anonymiserede CSV genereres fra datafilen.
5. **Testen** (gaten) genregner hele kæden fra råsvarene og fejler ved enhver afvigelse: en ændret celle, et tal i teksten uden dækning, en tankestreg, et scenarie uden sine forbehold, et landstal uden sin kilde.

Kommandoen `verify` i analysens værktøj kører præcis den kontrol: SHA256 af hvert råsvar mod listen, hver kolonne genudledt fra råsvaret, udvalget genregnet fra reserve-reglen, tal-blokken genregnet, den anonymiserede fil og scenarie-filen genregnet, og hver kilde til landets tal til stede i teksten.

02 Stikprøven

Populationen er defineret af et register, ikke af os: for parcelhuse er det Bygnings- og Boligregistret (BBR) med anvendelseskode 120 (fritliggende enfamiliehus), status opført.

Vægtning. Danmarks Statistiks boligopgørelse (BOL101: beboede boliger efter område og anvendelse, nyeste år) giver antallet af beboede parcel- og stuehuse i hver af de 98 kommuner. Hver kommune får en kvote efter største-rest-metoden (Hamiltons metode), så stikprøven ligger, hvor husene ligger; summen af kvoterne er stikprøvens størrelse.

Trækket i kommunen. Datafordelerens BBR-GraphQL sorterer en side efter bygningens tilfældige UUID, så den første side (1.000 bygninger) er en tilfældig delmængde af kommunens huse. Siden blandes med et frø pr. kommune, dubletter på samme adgangsadresse fjernes, registreringer med boligareal under 25 m2 fjernes (fejlregistreringer), og de første huse op til kvoten er primære; de næste (mindst tre, ellers tre femtedele af kvoten) er reserver i trækkets orden.

Reserve-reglen. Et svar tæller kun, når det er brugbart: beregningen lykkedes, motorens hovedbygning er et hus af den rigtige type (en adresse, hvis hovedbygning i motoren blev et udhus, tæller ikke), svaret er fra analysens systemmodel, og motoren fandt et hus at regne på (elforbrug og varmebehov begge over nul; et hus med opvarmet areal nul i registret giver et tomt svar). En primær adresse, der fejler eller er ubrugelig, erstattes af den næste reserve i samme kommune i trækkets orden; kvoten pr. kommune holdes præcis, og en anden kommunes hus bruges aldrig. Reglen er én funktion i værktøjet, og testen kontrollerer, at datafilens huse er præcis dem, reglen vælger ud fra svarlageret.

Kontrol af repræsentativitet. Regionernes andel af stikprøven sammenlignes med Danmarks Statistiks andele; husenes alder (fire årsbånd) ligeså. Afvigelser på få procentpoint er forventelige; større afvigelser tyder på en skævhed i trækket.

03 Motoren, modellen og prisgrundlaget

Samme motor som ejendomsejeren. Hvert hus regnes gennem den offentlige beregningsvej (poll-protokollen) med to eksplicitte valg: systemmodellen og ejerformen. Alt andet er motorens standard: husets egne data (BBR, energimærke, Danmarks Højdemodel med tagflader og skygge, luftfoto, netselskabets tariffer), solen time for time i vejråret fra PVGIS på hver tagflade, elpriserne fra de seneste tolv hele måneder i husets prisområde lagt på vejrårets timer med rangbevarende afbildning, husholdningens elforbrug fordelt på vejrårets timer som målte danske husstande uden elvarme (Energi Data Service, Energinet og netselskaberne), det varme vand tappet efter EU's tappeprofil med et morgen- og et aftenbad, de nye anlægs strømforbrug i tomgang (inverteren om natten og batteriets styring, efter producenternes datablade og HTW Berlins målinger), og en blandet heltals-optimering over årets 8.760 timer. Ændres motoren efter en kørsel, regnes de huse, ændringen kan flytte, igen på den ændrede motor, og analysens rapport skriver resultatet i afsnittet om beregningen.

Systemmodellen i parcelhusanalysen er Standard, den model adressesøgningen åbner med: alle tagets egnede flader kan bruges, og overskudsstrøm, der sælges til nettet, tæller med i økonomien; batteri, varmepumpe og varmelager kommer kun med, når de betaler sig. Valget af model afgør, hvad analysen kan sige: modellen dimensionerer efter husets økonomi, så analysen måler et økonomisk potentiale for husstanden, ikke et udbud til nettet. Analysens første kørsel (11. og 12. oktober 2026) brugte Optimin, hvor solcellerne kun lå på tagets bedste flader og var dimensioneret til husets eget elforbrug.

To spørgsmål for et hus med fyr. Et hus med olie-, gas- eller brændefyr og vandbåren varme får regnet to hele svar: behold fyret (solceller, batteri, varmelager og solvarme ved siden af fyret) og skift til varmepumpe (hele omlægningen). Hvert svar regnes helt igennem med sin egen dimensionering og sine egne produkter; det svar med den højeste nutidsværdi til markedspris er husets svar, og når intet betaler sig, vises det bedste alternativ af de to. I analysen tæller et hus, hvor ingen af de to svar betaler sig, som "byg intet".

Vindmøllen er enten ingen eller mindst den mindste rigtige mølle (1 kW), fordi ingen installatør sælger en mindre; over den grænse vælges størrelsen frit.

Prisgrundlaget er nøglefærdige anlæg: leverandørkatalogets listepriser inkl. moms, arbejdsløn og montage samt tilmelding, uden tilskud og uden rabat, regnet for en privat ejer. Varer fra andre webshops end hovedleverandøren koster butikkens pris plus halvdelen af hovedleverandørens listeavance, og hver af de øvrige leverandører koster 99 euro i fragt pr. projekt. Det er et konservativt grundlag: listepriser er ikke tilbudspriser, så tilbud under listepris gør økonomien bedre, og anlægget kan i mange tilfælde bygges billigere, end tallene viser. Et hus, hvor intet anlæg slår at lade være over 20 år, tæller som "byg intet".

Definitioner, der skal holdes konstante:

- "Betaler sig" = positiv nutidsværdi over 20 år med udskiftninger, restværdi og real kalkulationsrente = motoren bygger noget. Det er den samme mængde huse.
- "Tjent hjem inden for N år" = den simple tilbagebetaling (investering delt med første års besparelse), som motoren viser den med én decimal, er højst N år. Ved grænsen tæller husets viste tal: et hus med 10,04 år vises som 10,0 og tæller med inden for 10 år. Forskellen er få huse ud af tusind og ændrer ikke de afrundede andele.
- Medianerne (investering, besparelse, tilbagebetaling, nutidsværdi) er hver sin median over husene med anlæg, ikke ét hus.
- Varmepumpens effekt er termisk afgivet effekt; batteriets størrelse følger katalogets pakker af hybridinverter og batteri.
- "Solceller" er husets nye solceller på taget og på jorden; tagets egne tal (pladsen på taget, om anlægget fylder taget) gælder taget alene.

04 Systemdata pr. hus

Fra hvert råsvars energibalance udledes årsenergier for huset før og efter: solproduktion fra det nye anlæg (husets produktion efter minus før, så et anlæg i forvejen ikke tælles med), eget forbrug af solstrøm (produktion minus salg minus afkortet), salg til nettet, batteriets op- og afladning, den nye varmepumpes el inkl. dens elpatron, og køb fra nettet før og efter. Fra timeserierne udledes købet i aftentimerne, det største time-træk fra nettet, vinterkøbet og batteriets opladningstimer. Alle disse kolonner genudledes af testen fra råsvarene.

05 Skalering til landet og usikkerhed

Simpel skalering. Hvert hus i stikprøven står for N/n huse, hvor N er landets antal beboede parcel- og stuehuse (Danmarks Statistik) og n stikprøvens størrelse. Et landstal er stikprøvens sum for de huse, der indgår i scenariet, gange skalaen.

Kommunevægte som kontrol. Den implicitte vægt for et hus i kommune k er kommunens parcelhuse delt med dens kvote. Begge metoder regnes; afvigelsen rapporteres og skal være lille (brøkdele af en procent for de store scenarier).

Bootstrap-interval. Usikkerheden er et 95 %-interval fra 2.000 gentræk af stikprøven med tilbagelægning med fast frø, så intervallet er deterministisk og testen kan genregne det. For andele gives desuden den binomiale normalapproksimation. Stikprøven er kvoteret på kommune, ikke simpelt tilfældig, så intervallerne er vejledende.

Scenarier (adoptionsniveauer). Et scenarie er en delmængde af stikprøvens huse, skaleret til landet: husene med tilbagebetaling inden for 8 år, inden for 10 år, og alle huse, hvor et anlæg betaler sig over 20 år (= alle huse, motoren bygger). For hvert niveau opgøres huse, ny solcelleeffekt (tag og jord), de nye anlægs produktion, ændringen i eget forbrug og salg, batterikapacitet og afladning, varmepumper (antal, termisk effekt og elforbruget i husene med varmepumpe), køb fra nettet før og efter, investering, besparelse og CO_2, hver med sit bootstrap-interval. Scenariernes produktion sammenlignes på årsbasis med landets elforbrug, dagens solcelleproduktion og den registrerede solcelleeffekt.

Landets tal hentes af værktøjet fra Energinets åbne data (Energi Data Service: registreret solcelleeffekt pr. kommune, nyeste måned; årets målte produktion, forbrug og udveksling) og suppleres med Energistyrelsens energistatistik og Energinets redegørelse for elforsyningssikkerhed. Hvert tal gemmes med kilde, url og dato, og testen kræver, at hvert tal står i rapporten med sin kildelinje.

06 Hvad tallene kan bære, og hvad de ikke kan (regnskabet)

Analysen er en én-hus-model uden samtidighed, med perfekt fremsyn, ét vejrår, dagens priser holdt flade, skønnet forbrug (intet hus er regnet med et målt forbrug), nøglefærdige anlæg til listepris og uden tilskud. Det afgør, hvad der må siges:

Kan siges: andelen af huse, hvor et anlæg betaler sig, og hvor det er tjent hjem inden for N år; medianerne hus for hus; fordelinger på region, alder og opvarmning med deres intervaller; de nye anlægs årsproduktion og husenes eget forbrug og køb fra nettet før og efter som årsenergi; skalerede landstal som et økonomisk potentiale ved dagens priser med interval; sammenligningen med landets årsenergi; hvor længe et hus med batteri kunne køre uden strøm fra nettet, hvis anlægget var indrettet til nødstrøm (simuleret fra husets egne timer, højst den effekt en nødstrømsboks på 16 A giver).

Kan ikke siges: at husene aflaster elnettet, eller en aflastning i MW (husets største time-træk fra nettet stiger med et batteri, fordi det lader om natten; der er ingen samtidighed i modellen); at husene øger forsyningsudbuddet eller sænker Danmarks importbehov (modellen dimensionerer efter husets økonomi, ikke efter nettets behov; Energinets importbehov er effekt i knaphedstimer om vinteren, hvor solceller producerer lidt); at husene har nødstrøm, resiliens eller ø-drift (tilvalget var slået fra i beregningen, så nødstrømstallene viser, hvad batteriet kunne, ikke hvad anlægget er bygget til); hjælp i en vinterkrise; "så meget sparer du i praksis" (perfekt fremsyn er en øvre grænse); tal med tilskud; rigtige husstandes målte forbrug; et landsgennemsnit for installatørpriser; en færdig dimensionering.

Forbehold, der altid følger scenarierne: årlig energi er ikke effekt i knaphedstimerne; et batteri sælger kun strøm, huset selv har produceret (sol og vind), aldrig strøm, det har købt, og hæver husets spidstræk; varmepumperne er ny ellast, mest om vinteren; én-hus-model uden samtidighed; et økonomisk potentiale ved dagens priser for nøglefærdige anlæg til listepris, hverken prognose eller teknisk potentiale. Sætningerne står ordret i værktøjet, og testen kræver dem i rapporten, hver gang scenarierne vises.

07 Udgivelsesformatet

- **Analysesiden** (to døre, én side): ruten i appen og det forudrendrede dokument, samme komponent, alle tal fra den genererede JSON. Tabeller pr. region, alder og opvarmning, systemdata, scenarierne med forbehold og kilder, metode og stikprøven mod landet.
- **Den anonymiserede CSV** (den offentlige fil): ét hus pr. række uden adresse, adresse-id, postnummer, kommune og præcist opførelsesår; boligareal rundet til 10 m2 og loftet ved 300; rækkefølgen følger de ikke-identificerende kolonner. Den fulde fil med adresser er intern, og råsvarene (der bærer ejeroplysninger fra tinglysningen) ligger uden for repoet med en SHA256-liste, så hvert tal kan genregnes uden gensolve.
- **Pressesiden** med analysens boilerplate, citatform og skærmbilleder; pressematerialets tal står mellem markører og kontrolleres mod tal-blokken.
- **Tal-blokken** i analysens markdown er kilden til alle citater; teksten uden for blokken må ikke bære et procent- eller kronetal, der ikke står i blokken.

08 Stabilitetskontrol: en ny runde

En analyse står stærkere, når en ny, uafhængig stikprøve giver det samme. Opskriften: et nyt frø, mindst 200 nye huse trukket med samme regel (ingen af første rundes adresser), samme motor og samme prisvindue, samme udtræk og samme definitioner. Andelene og medianerne i runde 2 sammenlignes med runde 1: et tal holder, når runde 1's værdi ligger inden for runde 2's 95 %-interval, og forskellen mellem runderne ikke er statistisk sikker (dens eget interval dækker nul), så testen går begge veje. Ligger et tal uden for, undersøges årsagen (motorversion, prisvindue, trækket), og runde 2's rapport skriver den: motor-pariteten viser, om motoren er den samme, og omvægtningen til runde 1's fordeling på opvarmning viser, hvor meget af forskellen der er sammensætning. Et tal, der ikke holder, citeres kun med forbeholdet. Runde 2 får sit eget værktøj og sin egen rapport og rører ikke første rundes filer; første rundes tal ændres aldrig af en senere runde.

09 Opskrift til de næste analyser

Hver ny analyse følger de samme fem trin (stikprøve, beregning gennem den offentlige vej, udtræk gennem én dør, rapport fra datafilen, test der genregner alt) og skriver sine definitioner op, før den regner.

Etageejendomme

- **Første runde (oktober 2026):** private andelsboligforeninger i Københavns Kommune: etageboligbygninger (BBR 140), hvor Ejerfortegnelsen viser en privat andelsboligforening som ejer og BBR tæller mindst tre boliger; anlægget regnet på foreningens fælles el (trappelys, vaskeri, teknik) med foreningen som ejer. Analysen regnes igen på den gældende motor, før den udgives.
- **Population:** BBR anvendelseskoder 140 (etagebolig-bygning, flerfamiliehus) og evt. 130 (række-, kæde- og dobbelthus) som egen analyse; status opført. Enheden er ejendommen, ikke lejligheden: ejendommens bygninger med deres samlede boligareal og antal boliger (BBR enhedstal).
- **Vægtning:** Danmarks Statistik BOL101 med anvendelse etageboliger (og rækkehuse som egen kategori), beboede boliger pr. kommune; kvoter efter største-rest-metoden som i dag. Skaleringen til landet sker i antal bygninger, som kræver antal bygninger pr. kommune fra BBR (DST tæller boliger): tæl BBR-bygningerne pr. kommune ved udtrækket og gem tallet i udvalgsfilen.
- **Model og ejerform:** Standard med ejerformen efter Ejerfortegnelsen (andelsboligforening eller almen boligorganisation), så moms, afgift og skat følger ejerformen. Forbruget er ejendommens fælles el; fællesmåling af lejlighederne (ejendomsafregning) er en senere variant, der skal vælges eksplicit og siges højt.
- **Hvad motoren skal kunne:** ejendommens fælles elforbrug (det kan den i dag) og, til varianten med fællesmåling, lejlighedernes forbrug lagt sammen, varme typisk fjernvarme (ingen varmepumpe-historie i de fleste), taget er fladt eller lavt skrånende (tagformen fra højdemodellen skal håndtere flade tage med opstilling i rækker), og et større anlæg kan ramme netselskabets grænser. Ejerformens regler er i motoren i dag; fællesmålingen som antagelse skal kunne vælges eksplicit.
- **Hvad der kan siges:** andelen af ejendomme, hvor et taganlæg betaler sig på den fælles el, kWp pr. ejendom, besparelse pr. ejendom; besparelse pr. bolig kun med fællesmåling, og aldrig "pr. lejlighed i praksis".

Mark-solceller på fri jord

- **Population:** matrikler, ikke bygninger. Kandidater: landbrugsjord i landzone uden fredning, beskyttelseslinjer, Natura 2000, §3-natur, fredskov og kystnære zoner, og med en afstand til nærmeste netstation inden for motorens grænse. Datasæt: Matriklen (Geodatastyrelsen, Datafordeleren), Plandata (zonestatus, lokalplaner), Landbrugsstyrelsens markkort (IMK, hvilke marker er i omdrift), Danmarks Miljøportal (beskyttet natur), Energinets og netselskabernes kapacitetskort (tilslutningsmulighed).
- **Trækket:** matrikler pr. kommune med et seedet træk som i dag, kvoteret efter kommunens areal af landbrugsjord (Danmarks Statistik AFG07 eller IMK's markareal), mindst et par hektar sammenhængende.
- **Model og ejerform:** Optimax mark (feltet fyldt op til kapacitetsloftet) med ejerformen virksomhed eller landbrug (ex moms, afskrivning), prisperspektivet listepris; salg til nettet er her hele pointen, så Rc (systemvenlighed) og produktionsprofilen skal rapporteres.
- **Hvad motoren skal kunne:** et felt uden forbrug (ren eksport) skal kunne regnes med en nettilslutningspris, der følger afstanden til nettet og kapacitetskortet; producentbetalingen og tilslutningsreglerne skal være i prisgrundlaget; skyggen fra læhegn og terrænet fra højdemodellen skal regnes for marken, ikke taget.
- **Hvad der kan siges:** GWh pr. hektar, kr. pr. MWh produceret, andelen af marker med positiv nutidsværdi ved dagens priser; aldrig "forsyningssikkerhed" uden effekt i knaphedstimerne, aldrig et landstal uden nettilslutning som bindende forudsætning.

Fælles for begge

	PARCELHUSE (GJORT)	ETAGEEJENDOMME	MARK-SOLCELLER PÅ FRI JORD
Enhed	Bygning (BBR 120)	Bygning (BBR 140, evt. 130 som egen analyse)	Matrikel i landzone
Register til trækket	BBR via Datafordeleren	BBR via Datafordeleren	Matriklen, Plandata, IMK, Miljøportalen, kapacitetskort
Vægt pr. kommune	DST BOL101, beboede parcel- og stuehuse	DST BOL101, beboede etageboliger; BBR-bygninger pr. kommune til skaleringen	Landbrugsareal pr. kommune (DST AFG07 eller IMK)
Model	Standard (husets bedste økonomi)	Standard (ejendommens fælles el; fællesmåling som variant)	Optimax mark (feltet fyldt, salg til nettet)
Ejerform	Privat ejer	Andelsboligforening eller almen efter Ejerfortegnelsen	Virksomhed eller landbrug (ex moms)
Motoren skal kunne	Det den kan i dag	Den fælles el i dag; fællesmåling som eksplicit antagelse, flade tage i rækker	Felt uden forbrug med nettilslutningspris efter afstand og kapacitetskort
Må siges	Potentiale for husstanden, årsenergi	Potentiale på den fælles el pr. ejendom; pr. bolig kun ved fællesmåling	GWh pr. hektar, kr. pr. MWh, andel med positiv nutidsværdi
Må ikke siges	Aflastning af nettet, forsyningsudbud	Pr. lejlighed uden fællesmåling	Forsyningssikkerhed uden effekt i knaphedstimer, landstal uden nettilslutning

- Skriv tiers (hvem bygger) og sammenligningsgrundlag (landets tal med kilder) op, før runden regnes.
- Brug samme landstal-fil og samme kildeprincip (værktøjet henter, filen bærer kilde, url og dato).
- Kør en runde 2 med nyt frø, før tallene citeres som stabile.

10 Filer og kommandoer (parcelhusanalysen)

- `tools/datastudie_2026_10.py`: `sample` (stikprøven), `solve` (beregningen gennem backendens API), `report` (CSV, anonym CSV, SHA256-liste, scenarie-JSON, markdown; kræver råsvarene), `kilder` (landets tal fra Energinets åbne data og rapporttallene), `verify` (hele kæden genregnet; exit 1 ved drift).
- `docs/marketing/datastudie-2026-10.md` (tal-blokken), `-anonym.csv` (den offentlige fil), `-udvalg.json` (stikprøven), `-raasvar.sha256`, `-scenarier.json`, `-landstal.json`.
- `tools/datastudie-json.mjs` (analysesidens JSON fra CSV'en, scenarie- og landstal-filen), `tools/prerender-pages.mjs` (analysesiden, pressesiden og dette dokument som forudrendrede sider).
- Tests i `backend/app/engine/models/_tests.py`: `test_datastudie_2026_10`, `test_datastudie_2026_10_anonym_og_raasvar`, `test_datastudie_2026_10_landet_og_net`, `test_datastudie_2026_10_verify`, `test_datastudie_2026_10_scenarier`, `test_datastudie_page_json`, `test_datastudie_page_structure`, `test_presskit_figures`, `test_studie_metode_doc`.
