

## **Statistikdokumentation for Bogsalg 2023**

## 1 Indledning

Formålet med statistikken Bogsalg er at belyse udviklingen i styksalg af bøger i Danmark. Statistikken viser udviklingen i antal solgte eksemplarer fra fysiske bogforhandlere og websalg. Statistikken er ny fra 2024 og indeholder oplysninger fra tredje kvartal 2022 og frem.

## 2 Indhold

Bogsalg er en kvartalvis opgørelse af styksalget af fysiske bøger, e-bøger og lydbøger fra fysiske bogforhandlere (inkl. de største supermarkedskæder) og websalg. Salg af brugte bøger (fx fra antikvarier, genbrugsbutikker og loppemarkeder) er ikke inkluderet. Statistikken opgøres efter genre, format/medie, udgivelsessprog, originalsprog og indbindingstype. Statistikken formidles i Nyt fra Danmarks Statistik og i Statistikbanken. Datakilder og dækning er fortsat under udvikling, og derfor er tallene foreløbige. Fremtidige revideringer kan forekomme.

### 2.1 Indholdsbeskrivelse

Statistikken belyser udviklingen af styksalget af bøger i Danmark efter genre, format/medie, udgivelsessprog, originalsprog og typen af indbinding. Et salg svarer til et solgt eksemplar af en bog med et 13-cifret ISBN-nummer, der starter med 978 eller 979, og som er blevet solgt fra fysiske bogforhandlere og websalg. Hvis der fx ved et salg bliver solgt 10 eksemplarer af en bog, tæller det som 10 solgte bøger. Hver bog svarer til en unik udgave. Den samme titel har som regel et ISBN-nummer for hver udgave, og herunder har den fysiske udgave af en titel et andet ISBN-nummer end e-bogs- og lydbogsudgaven. Derudover kan andenudgaven af en given titel have et andet ISBN-nummer end den første.

Statistikken indeholder ikke på nuværende tidspunkt salg gennem streamingtjenester, dvs. bøger, der er købt adgang til via abonnement. Dette omfatter eksempelvis 30-dages abonnementsordninger, hvor man i perioden kan læse eller lytte til et ubegrænset antal bøger via en digital platform.

Statistikken vil løbende blive udvidet med flere kilder og indhold. Der vil blive arbejdet videre på at få etableret et robust datagrundlag om streaming af bøger, med henblik på at dække alle former for salg af bøger på markedet.

### 2.2 Grupperinger og klassifikationer

Genregrupperingen i denne statistik er dannet på baggrund af Themakoder, som er koder i det internationale klassifikationssystem: [THEMA](#), anvendt af den kommercielle bogbranche.

#### Hovedgenre

- Faglitteratur
- Skønlitteratur
- Børne- og ungelitteratur
- Uoplyst

#### Undergenre faglitteratur

- Biografier, sande fortællinger og ikke-fiktiv prosa
- Historie
- Kunst og kultur

- Mad og drikke
- Natur, miljø og geofag
- Personlig udvikling, relationer og sundhed
- Rejsebøger
- Religion og spiritualitet/filosofi
- Samfund, samfundsvidenskab og debat
- Økonomi & Business
- Hobby og livstil: øvrige
- Faglitteratur i øvrigt

### **Undergenre skønlitteratur**

- Digte
- Grafiske noveller
- Historiske romaner
- Krimier og Thrillere/spændingsromaner
- Romantik og Erotik
- Science fiction, fantasy o.l.
- Skønlitteratur i øvrigt

### **Undergenre børne- og ungelitteratur**

- Bøger for de helt små, billedbøger, aktivitetsbøger mv.
- Fagbøger til børn og unge
- Skønlitteratur for børn og unge
- Undervisningsmateriale til børn og unge
- Børn og unge øvrigt

Statistikken indeholder derudover oplysninger om medie/format, udgivelsesprog, originalsprog og indbindingstype.

### **Medie/format**

- Fysiske bøger
- E-bøger
- Lydbøger

### **Udgivelsesprog**

- Dansk
- Engelsk
- Svensk
- Norsk
- Tysk
- Fransk
- Øvrige eller flere sprog

### **Originalsprog**

- Dansk
- Engelsk
- Svensk
- Norsk

- Tysk
- Fransk
- Øvrige eller flere sprog
- Uoplyst

### **Indbinding**

- Indbundet, Hardback eller Hæftet
- Paperback og Pocketbøger
- Andet fysisk format (fx papbøger, spiralryg, anledningsalbums mv.)
- Uoplyst fysisk format
- Digitalt format

### **2.3 Sektordækning**

Kultur.

### **2.4 Begreber og definitioner**

ISBN: Står for Internationalt Standardbognummer. Et 13-cifret nummer som identificerer en bog, eller en udgave af en bog eller anden monografisk publikation fra et bestemt forlag. Ved indførelsen i 1970 og indtil 1. januar 2007 var numrene 10-cifrede.

### **2.5 Enheder**

Styksalg i antal solgte bogeksemplarer.

### **2.6 Population**

Populationen er bøger, som er solgt hos bogforhandlere enten i fysiske butikker eller som websalg.

### **2.7 Geografisk dækning**

Danmark

### **2.8 Tidsperiode**

Statistikken dækker tidsperioden fra 3. kvartal 2022 og frem.

### **2.9 Basisperiode**

Ikke relevant for denne statistik.

### **2.10 Måleenhed**

Antal solgte bogeksemplarer.

### **2.11 Referencetid**

Kvartal.

### **2.12 Hyppighed**

Årligt.

### **2.13 Indsamlingshjemmel og EU regulering**

Der er ingen EU-regulering for statistik om bogproduktion.

### **2.14 Indberetningsbyrde**

Statistikken er baseret på administrative registre. Der er derfor ingen direkte indberetningsbyrde i forbindelse med opgørelsen af denne statistik.

### **2.15 Øvrige oplysninger**

Øvrige oplysninger kan fås ved henvendelse til Danmarks Statistik.

## **3 Statistisk behandling**

Data til denne statistik indsamles via månedlige overførsler af transaktionsdata om styksalg af bøger fra SAXO.com A/S, Indeks Retail (Bog & Idé), Gucca og for de store supermarkedskæder (COOP, Salling og Dagrofa).

Transaktionsdata beriges med information fra DBK, som administrerer Bogportalen og Publizon, således at bøgerne kan klassificeres efter genre, format/medie, udgivelsessprog, originalsprog og indbindingstype. Såfremt disse oplysninger ikke findes for enkelte udgivelser søges disse udfyldt manuelt ved hjælp af forskellige antagelser.

### **3.1 Kilder**

Statistikken er baseret på transaktionsdata, som bliver leveret fra SAXO.com A/S, Indeks Retail (Bog & Idé) og Gucca. Hertil bliver der leveret transaktionsdata (stregkodedata) fra kontoret Priser og Forbrug i afdelingen Økonomisk statistik i Danmarks Statistik. Data indeholder alt styksalg af bøger i butikker under COOP, Salling og Dagrofa. Transaktionsdata indeholder bl.a. ISBN-nummeret for den bog, der er blevet solgt, tidspunktet for salget og antallet af solgte eksemplarer ved salgstidspunktet, samt et par enkelte supplerende oplysninger om den bog, som er blevet købt - fx om det er en fysisk bog, e- eller lydbog eller en anden form for vareoplysning, afhængig af dataleverandør.

ISBN-nummeret bruges som nøgle til at berige transaktionsdata med metadata fra DBK (Bogportalen) og Publizon. Oplysningerne i metadata anvendes til at klassificere bøgerne efter genre, format/medie, udgivelsessprog, originalsprog og indbindingstype.

### 3.2 Indsamlingshyppighed

Månedlig dataoverførsel fra dataleverandører.

### 3.3 Indsamlingsmetode

System-til-system-løsning.

### 3.4 Datavalidering

Etableringen af grupperingerne og klassifikationerne ud fra metadata er valideret fra essentielle aktører i bogbranchen og sagkyndige fagpersoner. Derudover sammenlignes data fra kvartal til kvartal. Ved store udsving kontaktes dataleverandør eller metadataleverandør.

### 3.5 Databehandling

Transaktionsdata renses for produkter, som ikke er bøger, fx Printpapir. Derudover fjernes ISBN-numre, der afviger fra den 13-cifrede struktur, som svarer til Bøger. Transaktionsdataene beriges med metadata fra DBK og Publizon. Metadataoplysningerne anvendes efterfølgende til at kategorisere data efter de klassifikationer, der ligger til grund for statistikbanktabellerne. Efter at data er beriget og færdigbehandlet aggregeres data i forbindelse med dannelse af tabeller til Statistikbanken.

Metadata indeholder bl.a. oplysninger om Themakoder, udgivelsessprog, originalsprog, format/medie og indbindingstype (digitalt format for e- og lydbøger). Themakoder er koder i et internationalt klassifikationssystem: THEMA, hvor bøger kategoriseres efter emner. Siden september 2019 er det blevet obligatorisk at klassificere i Thema i Bogportalen og ligeså i Publizon. Ikke alle udgivelser har fået tildelt themakoder på bogportalen, men dækningsgraden bliver højere i takt med, at oplysningerne på Bogportalen bliver opdateret.

Der foretages følgende manuelle rettelser:

#### Genre

Hvis der mangler Themakoder i metadata anvendes bogbutikkernes varegruppeoplysninger i metadata, som bliver brugt til manuel tildeling af Themakode. Dette sker ud fra fordelingen af Themakoder for titler i den samme varegruppe, hvis denne er sigende (fx Krimi og spænding, Fantasy, Skønlitteratur, Økonomi, Kogebøger, Skole- og lærebøger, Fagbøger for børn etc.). Suppleringen af Themakoder ud fra bogbutikkernes varegruppeoplysninger udgør cirka 1 procent af af salget i alt.

Hvis der mangler metadata om en bog, der er blevet solgt i et supermarked, suppleres der med oplysninger i transaktionsdata fra supermarkederne om de varer, der er blevet solgt. Hvis der under deres vareoplysninger, fx er anført "Krimier/Thrillere", så er denne information blevet brugt til at placere bogen i en genre. Denne genreplacering udgør omkring 3 procent af salget i alt.

Læs mere om metoden i dokumentet [Genreopdeling ved hjælp af Themakoder](#).

#### Udgivelsessprog

Hvis oplysninger om sprog findes i metadata fra DBK anvendes dette. Der kigges på de direkte oplysninger vedr. sprog, men også bogbutikkernes varegruppeoplysninger (varegruppe-tekst) i DBK-data, som også kan indeholde oplysninger om bogens udgivelsessprog. Hvis der ikke findes

oplysninger i metadata fra DBK eller Publizon, kigges der på udgivelsesland ud fra ISBN-nummeret, og så antages det, at fx bøger, der er blevet udgivet i England er på engelsk, bøger udgivet i Frankrig er på fransk osv.

### **Format/medie**

Databehandlingen tager udgangspunkt i, at de solgte bøger primært er fysiske bøger. For at sikre korrekt kategorisering af formatet bliver ISBN-numrene valideret ved at sammenligne med metadata fra DBK og Publizon. Hvis metadata indikerer, at der er tale om e-bøger eller lydbøger, bliver formatet rettet tilsvarende. Publizon-data anvendes som en pålidelig kilde til dette, da Publizon er Danmarks største distributør af e-bøger og lydbøger. Derudover inddrages information fra SAXO og Indeks Retail for yderligere verificering. I SAXOs transaktionsdata angiver feltet "TYPE," om produktet er en fysisk bog, e-bog eller lydbog, mens Indeks Retail opdeler deres data i kategorierne "Bøger" og "Lydbøger."

### **Indbindingstype**

I metadata fra DBK er de direkte oplysninger vedr. indbindingstype mangelfulde, derfor tjekkes også bogbutikkernes varegruppeoplysninger i data (varegruppe-tekst), som også kan indeholde oplysninger om bogens indbinding.

## **3.6 Korrektion**

Der foretages ikke korrektioner, ud over hvad der allerede er beskrevet under 3.4 Datavalidering og 3.5 Databehandling.

## **4 Relevans**

De centrale brugere er aktører i bogbranchen, Kulturministeriet og andre interesserede. Statistikken bliver brugt til analyser på litteraturområdet.

### **4.1 Brugerbehov**

De centrale brugere er aktører i bogbranchen, Kulturministeriet og andre interesserede. Statistikken bliver brugt til analyser på litteraturområdet.

### **4.2 Brugertilfredshed**

Statistikken er ny, men foreløbige og overordnede resultater fra tidligere perioder end de, der er offentliggjort i Statistikbanken er blevet valideret og godkendt af aktører i bogbranchen og Kulturministeriet.

### **4.3 Fuldstændighed af data**

Ikke relevant for denne statistik.

## 5 Præcision og pålidelighed

Statistikken dækker hovedparten af styksalget af bøger i Danmark, købt fra ny. Metadatadækningen vurderes tilfredsstillende for danske udgivelser, dog er dækningen lavere for udenlandske udgivelser. Dækningen er også lavere for bøger/bogprodukter, hvor formålet ikke er læsning, fx malebøger, pegebøger og lignende aktivitetsbøger, kalligrafi, bøger med musiknoder, sangbøger og bøger med Sudoku. ISBN-fejl kan forekomme, men vurderes ikke at være et generelt problem.

Tallene er foreløbige og fremtidige revisioner kan derfor forekomme i takt med udvidelse af bogsalgsdata og kvalitetsforbedringer af metadata.

### 5.1 Samlet præcision

Statistikken dækker hovedparten af styksalget af danske bøger, men ikke alt. Fx dækkes ikke salg af brugte bøger (antikvariatere, genbrugsbutikker, loppemarkeder) eller små selvstændige boghandlere og selvudgivelser på nettet.

Det kan forekomme, at et ISBN-nummer er blevet registreret forkert, enten i transaktionsdata eller på Bogportalen og/eller Publizon. Det kan potentielt bidrage til, at bøger placeres forkert i statistikken, men det vurderes ikke være et problem af generel karakter.

### 5.2 Stikprøveusikkerhed

Ikke relevant for denne statistik.

### 5.3 Anden usikkerhed

Statistikken dækker ikke salg af brugte bøger, salg fra små selvstændige boghandlere eller selvudgivelser på nettet. Derudover er metadatadækningen lavere for udenlandske udgivelser og visse typer hobbybøger som aktivitetsbøger og musiknoder.

Statistikken dækker heller ikke salg gennem streamingtjenester, dvs. bøger købt via abonnement.

Statistikken vil løbende blive udvidet med flere kilder og indhold. Der vil blive arbejdet videre på at få etableret et robust datagrundlag om streaming af bøger, med henblik på at dække alle former for salg af bøger på markedet.

### 5.4 Kvalitetsstyring

Danmarks Statistik følger anbefalinger vedrørende organisering og styring af kvalitet, der er givet i Adfærdskodeks for europæiske statistikker (Code of Practice, CoP) og den tilhørende implementeringsmodel Quality Assurance Framework (QAF). Læs mere om disse på [Adfærdskodeks for europæiske statistikker](#). Der er etableret en arbejdsgruppe for kvalitet og en central kvalitetssikringsfunktion, der løbende gennemfører tjek af produkter og processer.



## 5.5 Kvalitetssikring

Danmarks Statistik følger principperne i Adfærdskodeks for europæiske statistikker (Code of Practice, CoP) og bruger den tilhørende implementeringsmodel Quality Assurance Framework (QAF) ved implementeringen af disse principper. Dette indebærer løbende decentrale og centrale tjek af produkter og processer på baggrund af dokumentation, der følger internationale standarder. Den centrale kvalitetssikringsfunktion rapporterer til arbejdsgruppen for Kvalitet. Rapporteringen indeholder blandt andet forslag til forbedringer, som vurderes, besluttet og implementeres.

## 5.6 Kvalitetsvurdering

Statistikken er baseret på data fra store bogforhandlere og supermarkedskæder samt metadata fra de to ledende distributionsplatforme i Danmark. Dette sikrer en høj dækning af styksalget af bøger i Danmark.

Statistikken dækker dog ikke salg af brugte bøger, salg fra små selvstændige boghandlere eller selvudgivelser på nettet.

Dækningen er tilfredsstillende for danske udgivelser, men lavere for udenlandske udgivelser og visse typer hobbybøger. Det blev i 2019 obligatorisk, at bogudgivelser skulle registreres med Themakoder på Bogportalen og Publizon. Det forventes derfor, at kvaliteten af metadata for bøger, der er blevet udgivet fra 2019 og frem, er højere end for udgivelser fra før 2019.

Fremtidige revisioner af tallene kan forekomme i takt med forbedringer af metadata og udvidelse af datakilder.

## 5.7 Revisionspolitik

Danmarks Statistik foretager revisioner i offentliggjorte tal i overensstemmelse med [Danmarks Statistiks revisionspolitik](#). De fælles procedurer og principper i revisionspolitikken er for nogle statistikker suppleret med en specifik revisionspraksis.

## 5.8 Praksis for revisioner

Tallene er foreløbige, og fremtidige revisioner kan forekomme.

## 6 Aktualitet og punktlighed

Tallene er foreløbige, og fremtidige revisioner kan forekomme. Offentliggørelsen af statistikken sker ultimo november. Statistikken er ny, hvorfor statistikkens punktlighed endnu ikke kan beskrives.

### 6.1 Udgivelsestid for foreløbige og endelige tal

Tallene er foreløbige, og fremtidige revisioner kan forekomme. Offentliggørelsen af statistikken sker ultimo november.

### 6.2 Publikationspunktlighed

Statistikken er ny, hvorfor statistikkens punktlighed endnu ikke kan beskrives.

## 7 Sammenlignelighed

Statistikken er ikke direkte sammenlignelig med andre udgivne statistikker om bogsalg. Der findes ingen internationale retningslinjer for udarbejdelse af bogsalgsstatistik.

### 7.1 International sammenlignelighed

Der findes ingen internationale retningslinjer for udarbejdelse af bogsalgsstatistik.

### 7.2 Sammenlignelighed over tid

Statistikken er ny fra 2024 og indeholder oplysninger fra tredje kvartal 2022. Statistikken er sammenlignelig i hele perioden.

### 7.3 Sammenhæng med anden statistik

Der findes ikke direkte sammenhæng med anden statistik.

### 7.4 Intern konsistens

Data er baseret på transaktionsdata fra de store bogforhandlere, boghandelskæder og de store supermarkeds-kæder. Der vurderes, at indholdet i data fra boghandlere og supermarkederne er konsistente på tværs af datakilder. Indholdet i metadata fra DBK og Publizon vurderes ligeså at være sammenligneligt, da begge kilder indeholder de relevante oplysninger, der anvendes i udarbejdelsen af statistikken.

## 8 Tilgængelighed

Statistikken udgives i [Nyt fra Danmarks Statistik](#). I Statistikbanken offentliggøres tal for Bogsalg under emnet [Litteratur](#).

### 8.1 Udgivelseskalender

Udgivelsestidspunktet fremgår af udgivelseskalenderen. Datoen bekræftes i ugerne forinden.

### 8.3 Udgivelsespolitik - brugeroplysning

Statistikker offentliggøres altid kl. 08:00 på dagen, der er annonceret i udgivelseskalender. Ingen uden for Danmarks Statistik ser statistikken før offentliggørelsestidspunktet.

### 8.2 Udgivelseskalender - adgang

Udgivelseskalenderen kan findes på følgende link: [Udgivelseskalender](#).

### 8.4 NYT/Pressemeddelelse

Statistikken udgives i [Nyt fra Danmarks Statistik](#).

## 8.5 Publikationer

Ikke relevant for denne statistik.

## 8.6 Statistikbanken

Statistikken offentliggøres i Statistikbanken under [Litteratur](#) i følgende tabeller:

BOGSo1: Solgte bøger efter genre og format BOGSo2: Solgte bøger efter hovedgenre og udgivelsessprog BOGSo3: Solgte bøger efter genre, original sprog og udgivelsessprog

## 8.7 Adgang til mikrodata

Der er ikke adgang til statistikkens mikrodata.

## 8.8 Anden tilgængelighed

Ikke relevant for denne statistik.

## 8.9 Diskretioneringspolitik

[Datafortrolighedspolitik](#) i Danmarks Statistik følges.

## 8.10 Diskretionering og databehandling

Ikke relevant for denne statistik.

## 8.11 Reference til metodedokumenter

[Genreopdeling ved hjælp af Themakoder](#)

## 8.12 Dokumentation af kvalitetssikring

Resultater fra vurdering af beskrivelse af produkter og udvalgte processer foreligger i detaljeret form for hver statistik samt summarisk i rapporter til arbejdsgruppen for kvalitet.

## 9 Administrative oplysninger

Administrativt er statistikken placeret i kontoret Forskning, Teknologi og Kultur. Kontaktpersonen for statistikken er , tlf: og e-mail: .