

**Statistikdokumentation for  
Anløbsaktiviteten i danske havne 2020 (eksperimentel statistik)**

## 1 Indledning

Denne eksperimentelle statistik viser aktiviteten i danske havne målet ud fra skibsanløb. Statistikken er som eksperimentel statistik ikke en del af den officielle statistik. Statistikken betragtes som eksperimentel, fordi det er første gang Danmarks Statistik benytter de store mængder data, der ligger i AIS-systemet (Automated Information System - skibspositioner) og metoderne til behandling fortsat er under udvikling. Detaljeringen i statistikken og måden den formidles på vil også ændres hen ad vejen på baggrund af tilbagemelding fra brugere og den videre udvikling.

## 2 Indhold

Statistikken er et indeks over antallet af anløb i danske havne og dækker fragt- og containerskibe. Statistikken udarbejdes på grundlag af skibspositionsdata (AIS), som stilles til rådighed af Søfartsstyrelsen. Statistikken offentliggøres udelukkende i Statistikbanken.

### 2.1 Indholdsbeskrivelse

Statistikken opgør fragt- og containerskibes anløb i danske havne. Statistikken er en eksperimentel statistik, som baserer sig på AIS data, der indsamles og formidles af Søfartsstyrelsen.

### 2.2 Grupperinger og klassifikationer

I øjeblikket anvendes ingen grupperinger og klassifikationer

### 2.3 Sektordækning

Danske havne

### 2.4 Begreber og definitioner

AIS: AIS står for Automated Information System og er et internationalt sikkerhedssystem, hvor skibene udsender oplysninger om bl.a. position, retning og hastighed med henblik på at reducere risikoen for kollisioner. Faste søinstallationer såsom fyr og havvindmøller kan ligeledes udsende AIS meddelelser. Oplysningerne opsamles i Danmark af Søfartsstyrelsen til overvågning af skibstrafikken i de danske farvande.

Anløb: Et anløb er, når et skib ankommer til en havn. Anløbstidspunktet bruges til periodisering af anløbet.

### 2.5 Enheder

Statistikken offentliggøres på den geografiske enhed danske havne, som kan være en selvstændig juridisk enhed, flere juridiske enheder eller et eller flere arbejdssteder.

### 2.6 Population

Anløb af fragt- og containerskibe i danske havne.

## **2.7 Geografisk dækning**

Danmark

## **2.8 Tidsperiode**

2017-

## **2.9 Basisperiode**

2018 (gennemsnit) = 100

## **2.10 Måleenhed**

Statistikken opgøres på antal anløb.

## **2.11 Referencetid**

Statistikken opgøres på kalendermåned.

## **2.12 Hyppighed**

Statistikken er månedlig.

## **2.13 Indsamlingshjemmel og EU regulering**

Statistikken opgøres som en eksperimentel statistik på baggrund af offentlige tilgængelige data fra Søfartsstyrelsen. Der er ikke en indsamlingshjemmel og statistikken udarbejdes som del af Danmarks Statistiks generelle opgave med at udvikle statistik, jf. lov om Danmarks Statistik, §1, stk. 2, litra 1.

## **2.14 Indberetningsbyrde**

Der er ingen indberetningsbyrde

## **2.15 Øvrige oplysninger**

Øvrige oplysninger kan fås ved henvendelse til Danmarks Statistik

### 3 Statistisk behandling

AIS-data fra Søfartsstyrelsen modtages løbende og registreres dagligt. Data gennemgår en datareduktion, hvor overflødige observationer fjernes for at lette den yderligere databehandling. Dernæst udvælges ankomstobservation og afgangobservation, hvor ankomst/afgang identificeres og sammenhørende par identificeres. Herefter afgrænses havne, ved hjælp af data om skibspositioner. Placering af ankomst/afgangsobservationer i havne og statistik, hvor ankomst/afgang knyttes til en konkret havn og statistikdatasæt dannes

Endelig foretages tabellering, indeksering og sæsonkorrektion.

#### 3.1 Kilder

Datagrundlaget er primært skibspositionsdata (AIS) indsamlet af Søfartsstyrelsen og stillet til rådighed som fællesoffentlige data. Danmarks Statistik har modtaget og lagret data siden marts 2016.

AIS-data indeholder skibspositioner mm. for skibe med AIS transponder, dvs. mindst alle skibe over 300 BT, fiskefartøjer over 15 meter samt alle passagerfærger. Blandt disse er udvalgt skibe registreret som fragt- eller containerskibe.

#### 3.2 Indsamlingshyppighed

Data modtages i et løbende stream fra Søfartsstyrelsen og registreres i interne databaser i Danmarks Statistik dagligt. Data udgives månedligt.

#### 3.3 Indsamlingsmetode

Data registreres elektronisk af Søfartsstyrelsen via VHF modtagere, der er placeret langs de danske kyster. Danmarks Statistik modtager løbende data fra Søfartsstyrelsen.

#### 3.4 Datavalidering

Der foretages i øjeblikket kun overordnet kontrol af mulige dataudfald i materialet, dvs. perioder hvor Danmarks Statistik af forskellige grunde ikke har modtaget data. Kortvarige dataudfald har kun marginal betydning, mens udfald på flere timer kan betyde, at anløb ikke registrere.

#### 3.5 Databehandling

Trinene i databehandlingen er 1. Datareduktion 2. Udvalgelse af ankomstobservation og afgangobservation 3. Afgrænsning af havne 4. Placering af ankomst/afgangsobservationer i havne og statistik

Det første trin er datareduktion. Formålet er udelukkende at reducere mængden af data, der skal behandles og består af tre dele: geografisk reduktion, reduktion i observationshyppighed og reduktion i analyserede skibstyper.

AIS data indeholder alle informationer fra AIS systemet, som de danske receivers modtager. Det inkluderer en lang række observationer, som ikke direkte vedrører Danmark, men snarere Tyskland, Norge eller Sverige. Metoden til geografisk afgrænsning er simpel, idet data afgrænses med en firkant, som dækker hele Danmark og dermed også den sydlige del af Sverige. De svenske havne

bliver frasorteret i senere trin. Her er der mulighed for raffinering med afgrænsning til observationer, der er inden for det danske søterritorium.

Skibene udsender signaler med et interval bestemt af hastighed og aktivitetstype samt type af senderen/transponderen. De hyppigste signaler kommer med 6 sekunders mellemrum. Når formålet er at undersøge anløb i havnene, er så hyppige data slet ikke nødvendige. Data reduceres derfor til den første observation per minut.

Aktiviteten i havnene kan dække over mange typer af aktiviteter. For at understøtte eksisterende havnestatistikker, som fokuserer på godshåndtering i havnene, er data reduceret til de skibstyper, der udelukkende anvendes til godsfragt: fragtskibe og containerskibe.

Det næste trin i processen er at identificere skibsanløbene med ankomst og afgang. Hvert skib efterlader en række af positionsoplysninger og målet er at identificere den observation, der repræsenterer ankomsten til en havn og den tilhørende observation, der repræsenterer afgang fra havnen.

Processen er principielt simpel: Datasættet sorteres efter skibsidentifikation og tid. Dernæst markeres alle observationer, hvor skibet går fra at sejle (markør sat til under way skifter til moored samtidig med at skibet går fra at bevæge sig (mere end 1 knob) til at lægge (næsten) stille). Det er potentielle ankomster. Det samme gøres for potentielle afgang, hvor skibet går fra at have markøren sat til moored til under way og det bevæger sig. Med alle potentielle ankomster og afgang matches de sammen i par. Langt hovedparten hænger fint sammen, men der er både ankomster med manglende afgang og afgang med manglende ankomster. Mulige forklaringer er:

- Lange havneophold: Den gennemsnitlige opholdstid i havnen er omkring 12 timer, men hvis den er flere dage lang, kan der i starten eller enden af den betragtede periode mangle den matchende observation. Det kompenseres delvis for ved at bruge data, der rækker ud over perioden, fx vente 5-6 dage før seneste måneds statistik udarbejdes. Meget lange ophold i havnen er sjældent forbundet med fragtsejlads, men snarere et behov for reparation eller andet og er formentlig ikke væsentlige i forhold til statistikkens formål.
- Slukket transponder: Det er (som regel) ikke ulovligt at slå transponderen fra, og man kan forestille sig, at skibe først får slået transponderen til, når havnen er anløbet eller får den slået fra i havnen og glemmer at tænde den igen, når havnen forlades.
- Dataudfald: Der findes udfald i data, hvor data fra en enkelt receiver ikke kommer ind i en periode, data ikke streames fra Søfartsstyrelsen eller hvor data ikke opsamles og lagres af Danmarks Statistik. Det sidste var særlig i opstartsfasen og udfaldene er blevet reduceret meget over tid.

Det tredje trin er at afgrænse anløb til de observationer, der rent faktisk er anløb i en havn. Der er på dette tidspunkt fortsat en lang række af ankomst/afgange, som ikke er i nærheden af en havn. Ydermere skal data afgrænses til danske havne. Der er fortsat observationer fra særligt svenske havne med i datagrundlaget, herunder aktiviteten i Nordens største godshavn, Göteborg. Hvis data fra starten bliver reduceret til dansk søterritorium, forsvinder denne sidste del også.

Når dette tredje trin er lavet på mindst et par års data, kan en enkelt måneds statistik køres uden at danne en ny havneafgrænsning. Det reducerer den månedlige produktionstid markant og indarbejdelsen af nye data i afgrænsningen kan foretages efter statistikproduktionen.

Selve processen i dette trin består af tre dele:

- Sammenknyt observationer (ankomster bruges) i grupper (clusters) efter den indbyrdes afstand mellem observationerne og antallet af observationer i nærheden.
- Dan polygoner, som omslutter alle observationer i samme cluster
- Forbind de enkelte polygoner til faktiske havne

Resultatet af de tre trin er en spatial opslagstabel, som består af polygoner og tilhørende havn. En havn kan bestå af flere polygoner, mens en polygon kun kan være tilknyttet en havn. Hvis en ankomst ligger inden for en given polygon, ved vi dermed hvilken havn, anløbet hører til. For større havne repræsenterer hver polygon typisk et særskilt kaj anlæg.

Første trin består i at beregne alle indbyrdes afstande mellem ankomstpunkterne, hvorefter ankomsterne samles i grupper. Hvis to observationer (uafhængigt af tid) er mindre end fx 50 meter fra hinanden, knyttes de sammen. Hvis en tredje observation er mindre end 50 meter fra en af de første to observationer, bliver den også en del af gruppen. Til endelige grupper bliver alle grupper med flere end fx 5 observationer. Antallet af observationer i en gruppe og afstanden mellem observationer i gruppen er parametre, der kan tilpasses. Jo færre observationer, der anvendes des større bør afstanden være og des mindre bør antallet af observationer i en gruppe være. Med tre års data anvendes 70 meter og 5 observationer. Efter første skridt har alle observationer fået tilknyttet en gruppe (cluster). Alle observationer, der ikke opfylder kriterierne får clusternummer -1 og betragtes som falsk positive anløb.

Andet trin danner polygoner, der omkranser de enkelte clusters, således at alle observationer i samme cluster ligger indeni eller på kanten af polygonen. Hvis man ser på beliggenheden af disse clusters, får man uden videre bearbejdning et godt billede af, hvor der er havne. I lande, hvor der etableres uofficielle havne, kan man her identificere uofficielle havne.

Tredje skridt knytter de enkelte polygoner til en havn. Det gøres i en iterativ proces, hvor der tages udgangspunkt i et havnecentroid (eller tilsvarende – i praksis anvendes oftest koordinatet for havnen, som angivet i UnLocode listen, men det kan også findes ved simpel opslag i fx Google Maps). Den iterative proces sikrer at ingen clusters knyttes til mere end en havn og at polygonen knyttes til den havn, der ligger tættest på polygonen. Tætliggende havne, som skal kunne adskilles, kan give lidt manuelt arbejde med at tilpasse udgangspunkterne, så de enkelte cluster kommer med i den rette havn. Iterationen foregår ved, at der dannes en gradvis større cirkel omkring udgangspunktet for havnen. Hvis cirklen overlapper en clusterpolygon, tilknyttes denne cluster til den pågældende havn, hvorefter den ikke kan kobles til andre havne. Den maksimale afstand fra havnens udgangspunkt for clusters må defineres ud fra konkrete data. I de danske data er begrænsningen sat efter, at der ud for Skagen er defineret en række clusters for skibe, der ligger på red. De betragtes reelt ikke som værende i havn og grænsen er sat, således, at disse clusters ikke medtages. Da der udelukkende tages udgangspunkt i danske havne, sorterer processen også de primært svenske havne fra.

Ved enden har vi en opslagstabel, hvor de øvrige oplysninger om konkrete havne også kan findes, fx kommune, unlocode, navn, CVR nummer og kystzone.

Det sidste trin i hele processen er, at anløbene (defineret som ankomst og tilhørende afgang i trin 2) matches op mod polygonerne, som definerer (dele af) havnene. Resultatet er et datasæt, som indeholder oplysninger om anløbet (primært tid, skibsidifikation og skibsnavn) og den havn, der er anløbet (primært havnens navn, kode og landsdel). Data kan yderligere beriges med mere detaljerede skibsoplysninger fra skibsregistre, fx størrelse, flagstat, mere præcis skibstype og ejer/operatør.

Herfra er processen helt standard for statistikproduktion. Tabellering på eventuelle på underopdelinger, indeksering og sæsonkorrektion.

### **3.6 Korrektion**

Der foretages sæsonkorrektion.

## 4 Relevans

Statistikken giver hurtigere data om aktiviteten i danske havne og mulighed for at belyse flere typer af fartøjer i forhold til de eksisterende skibsfartsstatistikker. Statistikken er eksperimentel og en del af dette koncept er at afdække relevans og brugerbehov. Brancheorganisationer har dog udtryk interesse for opgørelsen.

### 4.1 Brugerbehov

Søfartens brancheorganisationer har udtrykt interesse i opgørelsen. Anløbsindikatoren fordel er primært, at den kommer med tidligere tal og kan afdække alle skibe i dansk farvand i forhold til de eksisterende søfartsstatistikker. Processen omkring en eksperimentel statistik handler også om, at afdække om statistikken er relevant og i hvilken form.

### 4.2 Brugertilfredshed

Statistikken tilgodeser et brugerønske om hurtigere data om aktiviteten i danske havne og kan derfor supplere de eksisterende søfartsstatistikker, som kommer kvartalsvis.

### 4.3 Fuldstændighed af data

Der er ikke definerede krav til statistikken.

## 5 Præcision og pålidelighed

Statistikken præcision i forhold til udviklingen i de faktiske anløb i danske havne og for fragt- og containerskibe er relativ høj baseret på en sammenligning med den kvartalsvise Søfartsstatistik.

Statistikken vurderes for indeværende at have en mellem pålidelig, idet der endnu ikke er foretaget revisionsanalyser af statistikken. Datagrundlaget ændrer sig ikke over tid, så revisioner vil skyldes tilpasninger i metode.

### 5.1 Samlet præcision

Det samlede antal anløb for de største havne i Danmark kan sammenlignes med indberetninger til statistikken Skibsfart på større havne. Sammenligningen viser, at der er en gennemsnitlig bias på -3 pct. over de seneste tre år. Retningen af bias er i overensstemmelse med forventningerne givet mulige fejlkilder.

En sammenligning mellem indeks baseret på indberetninger til Skibsfart på større havne og aktivitetsindikatoren giver en gennemsnitlig bias på -0,8 indekspoint de seneste sammenlignelige 12 måneder og en gennemsnitlig absolut afvigelse på 1,1 indekspoint.

### 5.2 Stikprøveusikkerhed

Der er ikke tale om en stikprøve og punktet er således ikke relevant.

### 5.3 Anden usikkerhed

Der kan principielt forekomme dækningsfejl, målefejl og bortfald, som alle vil påvirke antallet af anløb negativt. Det er ikke umiddelbart klart, om det påvirker det publicerede indeks, men a priori forventes fejlene at være konstante over tid og derfor ikke at påvirke indekset, hvor alle anløbene vægtes ens. Et indeks baseret på bruttotonnage er mere følsomt overfor de potentielle fejl, da de enkelte anløbs vægt i indekset er forskellige.

I den videre udvikling af statistikken foretages en analyse af fejlkilderne og deres eventuelle påvirkning på resultaterne. De fejlkilder, der kan løses gennem bedre metoder, vil blive søgt elimineret.

### 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, beslattes og implementeres.

### 5.6 Kvalitetsvurdering

Der er ikke foretaget kvalitetsmåling endnu, men det vurderes ud fra sammenligning med eksisterende statistikker, at statistikken overordnet set har en god kvalitet.

### 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

Der foretages løbende revisioner af statistikken for hele den publicerede periode.

Revisioner skyldes dels nye data, som påvirker afgrænsningen af, hvad der indgår i den hidtidige statistik og dels af ændringer i metoder, som løbende evalueres.

Der er ikke foretaget revisionsberegninger endnu.



## 6 Aktualitet og punktlighed

Statistikken er planlagt til at blive offentliggjort ca. 10 dage efter udløbet af referenceperioden og der er ingen forventning om forsinkelser.

### 6.1 Udgivelsestid for foreløbige og endelige tal

Udgivelsestiden planlægges at være 8-12 dage efter referenceperiodens udløb.

### 6.2 Publikationspunktlighed

Der er ikke statistik på publikationspunktlighed endnu.

## 7 Sammenlignelighed

Kun få andre lande bruger AIS-data til belysning af skibsfarten, da statistikken fortsat er en eksperimentel statistik uden en fastlagt international metode. Der kan sammenlignes med tilsvarende opgørelser baseret på statistikken for større danske havne på trods en række forskelle.

### 7.1 International sammenlignelighed

De fleste lande med havne udarbejder statistik over aktiviteten i havnene. De statistikker er typisk udarbejdet som direkte spørgeskemaer, men meget få lande, fx [Storbritannien](#) har brugt AIS-data til hurtigere statistik for havnenes aktivitet.

### 7.2 Sammenlignelighed over tid

Der er kun få års data til rådighed. For den periode er data fuldt sammenlignelig.

### 7.3 Sammenhæng med anden statistik

Statistikken kan sammenlignes med statistikken Skibsfart i danske havne, som baseres på direkte indsamling af data fra de større danske havne.

Der er en definatorisk afvigelse fra Skibsfart i danske havne, hvor det ikke er skibets type, men det faktum, at der er losset eller lastet gods, der afgrænser skibene. Det betyder, at der er to modsatrettede forskelle: for det første inkluderer Skibsfart i danske havne skibe, som ikke er registreret som fragt- eller containerskibe, hvilket giver statistikken flere skibe end den AIS-baserede statistik og for det andet inkluderer den AIS-baserede statistik alle skibe uanset fragt, hvilket potentielt giver den AIS-baserede statistik flere skibe end Skibsfart i danske havne. Hvorvidt den ene eller anden effekt dominerer, afhænger af den enkelte havn og deres øvrige ydelser end fragthåndtering, men for alle havne samlet har havnene registreret flere skibe end AIS-statistikken viser, hvilket indikerer, at langt de fleste skibe, der anløber danske havne, gør det med formålet at losse eller laste.

### 7.4 Intern konsistens

Der er intern konsistens i data

## 8 Tilgængelighed

Statistikken publiceres i [Statistikbanken](#) og på emnesiden for [Havne](#)

### 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

Der er ingen særskilt NYT-udgivelse for denne statistik

### 8.5 Publikationer

Der er ikke publikationer, der indeholder statistikken.

### 8.6 Statistikbanken

I Statistikbanken publiceres tabellen

- [AIS1: Anløbsaktivitet i danske havne \(eksperimentel statistik\)](#)

### 8.7 Adgang til mikrodata

Der er ikke taget stilling til om og i givet fald i hvilket omfang mikrodata vil blive stillet til rådighed for forskere. Hvis et ønske og behov opstår, vil spørgsmålet blive behandlet.

### 8.8 Anden tilgængelighed

Ikke relevant

### 8.9 Diskretioneringspolitik

[Datafortrolighedspolitik](#) i Danmarks Statistik følges.

### **8.10 Diskretionering og databehandling**

Datagrundlaget er offentlig tilgængeligt for alle via Søfartsstyrelsen og der foretages derfor ikke diskretionering i Danmarks Statistik, idet dog enkelte observationer/skibe ikke vil blive offentliggjort.

### **8.11 Reference til metodedokumenter**

Ikke relevant for denne statistik.

### **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 Konjunkturstatistik. Den statistikansvarlige er Peter Ottosen, tlf. 3917 3025, e-mail [pot@dst.dk](mailto:pot@dst.dk)

### **9.1 Organisation**

Danmarks Statistik

### **9.2 Kontor, afdeling**

Konjunkturstatistik, Erhvervsstatistik

### **9.3 Kontaktpersonens navn**

Peter Ottosen

### **9.4 Kontaktpersonens funktion**

Statistikansvarlig

### **9.5 Adresse**

Sejrøgade 11, 2100 København Ø

### **9.6 E-mailadresse**

[pot@dst.dk](mailto:pot@dst.dk)

## **9.7 Telefonnummer**

39 17 30 25

## **9.8 Faxnummer**

N/A