

Asger Dalsgaard, Elaheh Peyman
Granov, Lone Alletorp Callard,
Louise Damgaard Terkildsen,
Mike Wedel Hansen, Mikkel Bjørn
Frederiksen, Sine Kjær Stage

ENDAGS TURISME

Kan teledata kvalificere viden om
endagsturismen – og dermed understøtte
en mere bæredygtig turisme i Danmark?



Foto: Klitmoeller, Mette Johnsen

INDHOLDSFORTEGNELSE

1 INTRODUKTION	3
1.1 Projektets datapartnere & -leverandører	6
2 BÆREDYGTIGHED	7
3 PRÆSENTATION AF CASES	9
3.1 Indre by i København	10
3.1.1 Validering af KPI'er	11
3.1.2 Samlet konklusion på casen	16
3.2 Egeskov Slot	17
3.2.1 Variable	18
3.2.2 Validering af KPI'er	20
3.2.3 Konklusion på Casen	29
3.3 Landal SeaWest	30
3.3.1 Variable	31
3.3.2 Validering af KPI'er	32
3.3.3 Case-konklusion	36
3.4 Rammstein Koncert i Aarhus	36
3.4.1 Variable	37
3.4.2 Validering af KPI'er	37
3.4.3 Case-konklusion	47
4 KONKLUSION	48
5 HVORDAN SKAL MAN ARBEJDE MED TELEDATA ("HOW-TO-GUIDE")	51
5.1 Teledataudbydere	52
5.1.1 Grundlæggende forskelle i datastrukturen	53
5.1.2 3's Gæstedefinitioner	54
5.2 Udvælgelses af case / vurdering af data	55
5.2.1 Område, Masterdækning & Forventet antal gæster	57
5.2.2 Baseline	57
5.2.3 Realistiske forventninger	57
5.3 Styrker og svagheder ved teledata	58
5.3.1 Ekstrapolering	58
5.3.2 "Beboer"-kategorien (3)	58
5.3.3 Overnattende gæster	59
5.3.4 Antal observationer	59
5.3.5 Faktiske tal – eller trends?	60
5.4 Opsamling	60

INTRODUKTION

Foto: Lønstrup, Niclas Jessen



Bæredygtighed i turisme og sikring af langsigtet konkurrencedygtighed kræver udvikling og implementering af nye måder at måle turismen på. I Danske Destinationers projekt, Bæredygtige Destinationer i Praksis, er målet at udvikle indikatorer, der kan implementeres, håndteres og opdateres, og som skal sikre stærkere beslutningsgrundlag til at forme fremtidens bæredygtige turisme- og destinationsudvikling, samt synliggøre effekten af destinationsselskabernes arbejde og indsats for at fremme dansk turismes bæredygtige værdiskabelse.

For at kunne sikre en bæredygtig og balanceret turismeudvikling på destinationerne er det nødvendigt at kende volumen af gæster og lokale, samt deres mobilitet, på de enkelte destinationer. Mens data om antallet af endagsturister kan sige noget om volumen og presset på destinationen overordnet, kan mobilitetsdata øge forståelsen af gæstens brug af destinationen lokalt og forklare, hvorfor nogle områder oplever flere og måske for mange gæster på samme tid og sted. Data om (endags)turisters bevægelse kan hjælpe til at udvikle lokale spredningsstrategier og sikre en mere bæredygtig turisme på sigt. Formålet med dette projekt i projektet, "Endagsturisme", er derfor at blive bedre i stand til at måle og forstå endagsturismen på de enkelte destinationer, og bedre kunne agere herpå.

Hidtil har endagsturismen både i Danmark og i de enkelte destinationer været meget dårligt belyst. Målemetoderne brugt i dansk turisme har kun i ringe grad kunnet måle endagsturismen, og data har været forældet og langt fra tilstrækkeligt. Dette projekt afsøger derfor mulighederne for bedre at kunne estimere mængden af endagsturister gennem teledata. Projektets indledende målsætninger er således;

- at hjælpe destinationerne med at validere og forstå data omkring endagsturisme lokalt
- at understøtte korrekt anvendelse af teledata i den kommende nationale dataplatform, og gøre destinationerne i bedre stand til at agere på evt. friktionspunkter, mm., samt
- at understøtte, at der med en række lokale valideringer af endagsturismens størrelse (antal endagsturister), vil være grobund for at skalere og forstå endagsturismens størrelse på nationalt niveau.

Målsætninger er forsøgt indfriet gennem valideringscases, som så vidt muligt, dækker forskellige scenarier for endagsturismen. I alt blev fem cases udvalgt:

1. en koncert
2. en attraktion
3. et feriehusområde
4. et byområde
5. endagsturisme til omkringliggende kommuner fra hovedstaden

De fem cases blev udvalgt på tre forskellige geografiske niveauer, i forsøget på bedst muligt at afdække styrker og svagheder i datatypen (teledata) i forskellige scenarier. Projektet har således arbejdet med fem cases på tre forskellige geografiske niveauer, som hver især har sin styrke til at bidrage med målinger og forståelse af endagsturismen.

De tre niveau er:

1. Lille afgrænset geografisk område
(Rammstein koncert og Egeskov Slot)
2. Større, afgrænset, geografisk område
(Landal SeaWest og omkringliggende feriehusområde)
3. Kommunalt
(Indre by i København og hovedstadens omkringliggende kommuner)

Alle cases validerer antallet af gæster i det udvalgte område, samt mobiliteten enten ud eller ind i området. I valideringen blev en række KPI'er (Key Performance Indicators), som anses som relevante måle- og sammenligningspunkter på tværs af cases, opstillet. KPI'erne siger i sig selv ikke nødvendigvis noget om antallet af endagsturister, men for at kunne konkludere på mængden af endagsturister, skal den totale andel af gæster valideres, da endagsturisterne er et produkt heraf. Sagt med andre ord: Hvis den samlede gæstemængde vurderes som fejlbehæftet, kan mængden af endagsturister altså ikke vurderes, da endagsturismen afhænger af et retvisende tal på den totale gæstemængde.



Alle valideringerne søger at svare på en eller flere af følgende KPI'er:

- Antal gæster i POI'en (Point of Interest) – og hvilke kommuner kommer de fra (antal + oprindelse)
- Antal gæster i POI'en, der rejser ud af POI'en under deres besøg (mobilitet + endagsturisme i omkringliggende kommuner)
- Antal gæster, der besøger X andre kommuner under et besøg (antal besøgs-kommuner)
- Antal gæster i POI'en i forhold til antal indbygger/km² (intensitet + densitet)
 - Dage / tidsperioder som ligger i 4-kvartil i forhold til antal gæster på destination (vurdering af potentiel max.-belastning)
- Opholdslængder (mobilitet)

Endeligt er det vigtigt at understrege, at rapporten og dens konklusioner hviler på fem cases, som på trods af at være udvalgt med tanke for at vise en bred repræsentation af teledatacases i turisme, ikke viser, hvad teledata kan og ikke kan i andre end de analyserede cases. Konklusionerne skal læses som retningsgivende ud fra de cases, der er arbejdet med i denne rapport – og ikke som entydige for alle typer cases.

1.1 PROJEKTETS DATAPARTNERE & -LEVERANDØRER

Projektet er udformet bl.a. med det formål at sammenligne data fra de to teledataudbydere Telia og 3. Det har kun kunne lade sig gøre, fordi både Telia og 3 aktivt søger et positivt samarbejde og dialog om deres data med de danske destinationsselskaber. De har desuden en stor interesse i at udvikle deres data og skabe transparens, så datakvaliteten kan blive endnu bedre. Telia og 3 har begge lagt mange timer i udvælgelsen af cases og i løbende møder og dialoger om projektets arbejde med projektdeltagerne.

Særligt tak til Oliva Lundholm og Claus Jelle fra Telia, samt Katla Hrund Björnsdottir og Thomas Stjernholm Pedersen fra 3.

BÆREDYGTIGHED

2



Bæredygtighed er allerede i dag – og forventes i endnu højere grad at blive – et konkurrenceparameter inden for turismebranchen. Både danske og internationale undersøgelser peger på øget efterspørgsel efter bæredygtighed fra forbrugere og turisterne som følge af COVID-19-krisen. De enkelte erhvervsaktører kan gøre meget på vegne af eget brand, men løftes i kraft af det samlede destinationsbrand og samlede fortælling omkring den bæredygtige destinationsoplevelse. Det er ikke nok, at de danske destinationer alene faciliterer erhvervets bæredygtige certificering eller tilbyder en enkelt guide blandt flere til bæredygtige oplevelser. Destinationerne skal samlet overbevise forbrugere om, at de er bæredygtige. Det kræver et markant skifte – fra bæredygtighed som et projekt, og som en tidsbegrænset øvelse, til bæredygtighed som en integreret praksis. Men de danske destinationer mangler seriøse og konkrete metoder til at måle turismens bæredygtige effekt og effekten af destinationsselskabernes arbejde for at fremme bæredygtig turisme- og destinationsudvikling.



Foto: Copenhagen Oerestad Mountain Dwellings facades

I Danmark taler vi om fremtidens turisme, men måler og belønner gårsdagens turismetilgang. Langt de fleste destinationer og destinationsstrategier anerkender vigtigheden af bæredygtig udvikling og balanceret vækst, men måles på klassisk turismevækst som overnatningstal og direkte omsætningstal.

Destinationerne spiller en afgørende rolle i realisering og implementering af nye nationale ambitioner for bæredygtig turismeudvikling. Den afgørende rolle har destinationerne brug for at kunne påvise – og derfor er der brug for konkrete, håndterbare og opdaterbare mål for den bæredygtige udvikling.

Dertil er der brug for, at den dybe tallerken ikke udvikles på 19 forskellige destinationer, men at vi sammen udvikler metoder og værktøjer til arbejdet med bæredygtighed, som kvalificeret beslutningsgrundlag til at drive den samlede nationale bæredygtige udvikling frem.

Dette projekt i "Bæredygtige Destinationer i Praksis" skal være et bidrag til at komme brugen af tilgængelige data nærmere. Vi forsøger at dykke ned i mål, som er meningsfulde i forhold til det arbejde og de bæredygtige ambitioner, de danske destinationer og deres omgivelser har – og hvordan vi kan måle disse på en måde, der ikke kræver større investeringer.

Samlet set skal cases beskrevet i denne rapport synliggøre destinationernes afgørende betydning for Danmarks bæredygtige turismeudvikling.

PRÆSENTATION AF CASES

3

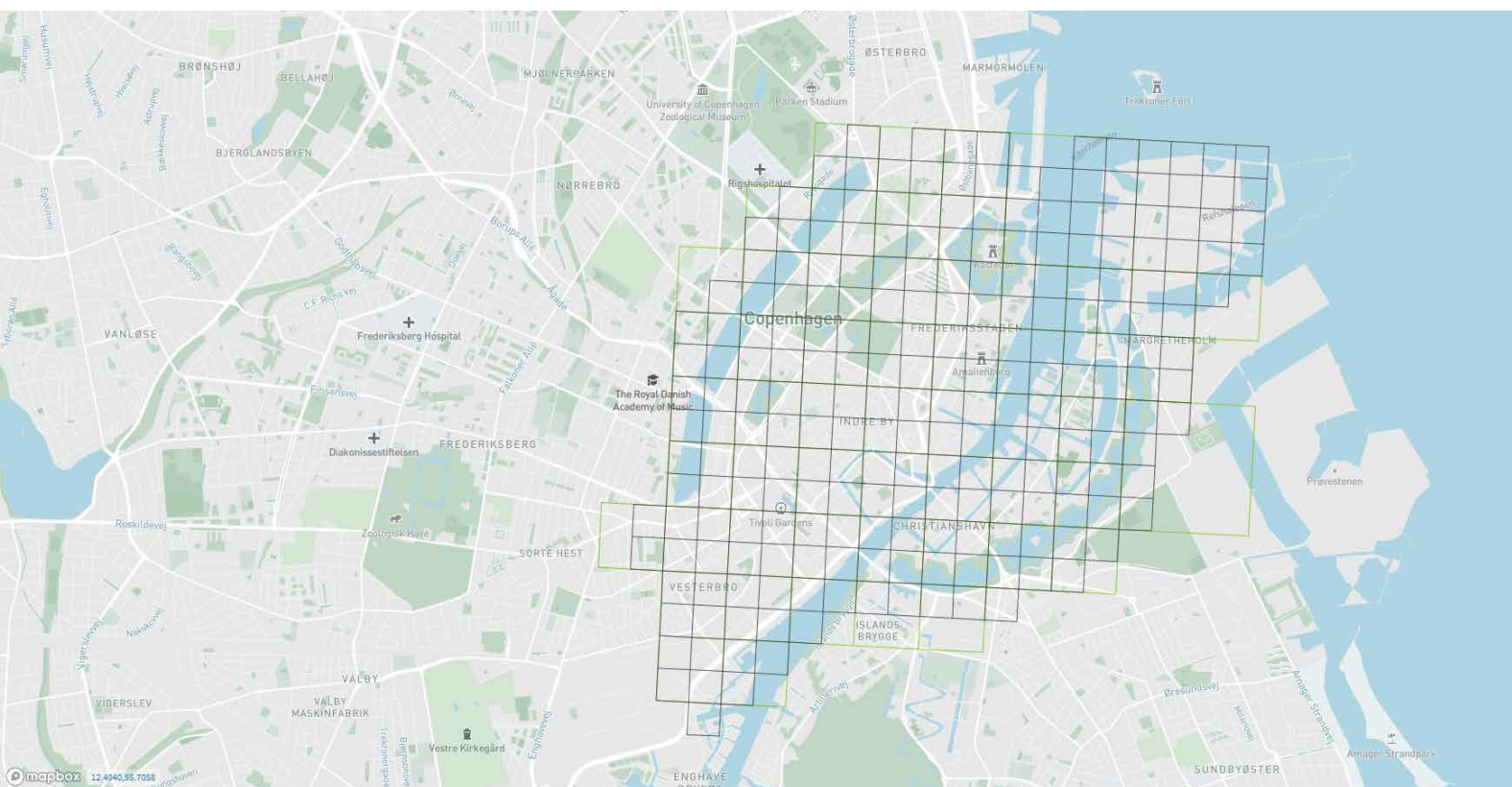


I følgende afsnit præsenteres fire ud af fem cases, der er arbejdet med i projektet. Casen om endagsturister i kommunerne omkring København er ikke medtaget, da det ikke har været muligt at finde data til at validere disse datasæt og det dermed ikke er muligt at lave en case for hvordan det skal bruges og hvad man skal tage højde for i den forbindelse. Det er ligeledes ikke været muligt at validere de to datasæt op imod hinanden grundet uregelmæssigheder i det ene datasæt.

3.1 INDRE BY I KØBENHAVN

Casen omhandler Københavns Indre By samt en lille del af det omkringliggende område (data dækker f.eks. også en smule af Vesterbro). Både 3 og Telia opdeler området i geografiske firkanter på baggrund af placeringen af telefonmaster i byen, men størrelsen af disse firkanter varierer på baggrund af udbyder. 3's data indeholder i alt 231 geografiske firkanter, som hver består af et 250*250 meter område, hvorimod Telias data indeholder 56 geografiske firkanter, som hver består af et 500*500 meter område.

Ser man dog på det samlede areal, så dækker begge datasæt 14 mio. kvadratmeter.



Begge datasæt indeholder data på timeniveau for juli og august måned 2022, med få undtagelser som resultat af begrænsninger i teleselskabernes evne til at videregive data på tværs af månedsskifter jf. den lovgivning, de er underlagt.

Datasættet fra 3 indeholder desuden data fordelt på nationalitet (danske / internationale besøgende) såvel som data fordelt på 6 besøgstyper; Beboere, i transit, pendler, endagsbesøgende og overnattende besøgende (kort og langt ophold). Datasættet fra Telia indeholder kun data for danskere fordelt på ophavskommune.

Målet med casen er at forsøge at validere data fra de to teleselskaber med henblik på at kunne bruge teledata til aktivt at følge med i besøgendes brug af forskellige områder i byen og den volumen af besøgende som er i byen på ethvert givent tidspunkt. Danmarks Statistik udgiver data på antallet af overnatninger i København by samlet og i de fire underliggende kommuner, men dette data udgives med to måneders forsinkelse og kan ikke brydes ned på mindre områder end kommune-niveau. Ønsket er derfor at kunne bruge teledata til at få et mere retvisende indblik i antallet og trykket af besøgende i byen såvel som at måle i hvilken grad de besøgende i byen spreder sig ud i byens forskellige bydele.

3.1.1 VALIDERING AF KPI'ER

Med henblik på at validere KPI'erne (besøgendes bevægelse rundt i København By og andele af besøgende i individuelle områder af byen), er disse sammenlignet med andre datakilder som bruges til at vurdere antallet af besøgende i København på dagsniveau:

Sensor data Reklamestandere mm. rundt om i København har WIFI-sensorer som måler antallet af personer som bevæger sig indenfor en given afstand af denne sensor. Aggregeret giver data fra sensorerne et billede af mere overordnede bevægelsesmønstre i byen, som kan bruges til at sammenligne med teledata som også måler overordnede bevægelsesmønstre i byen

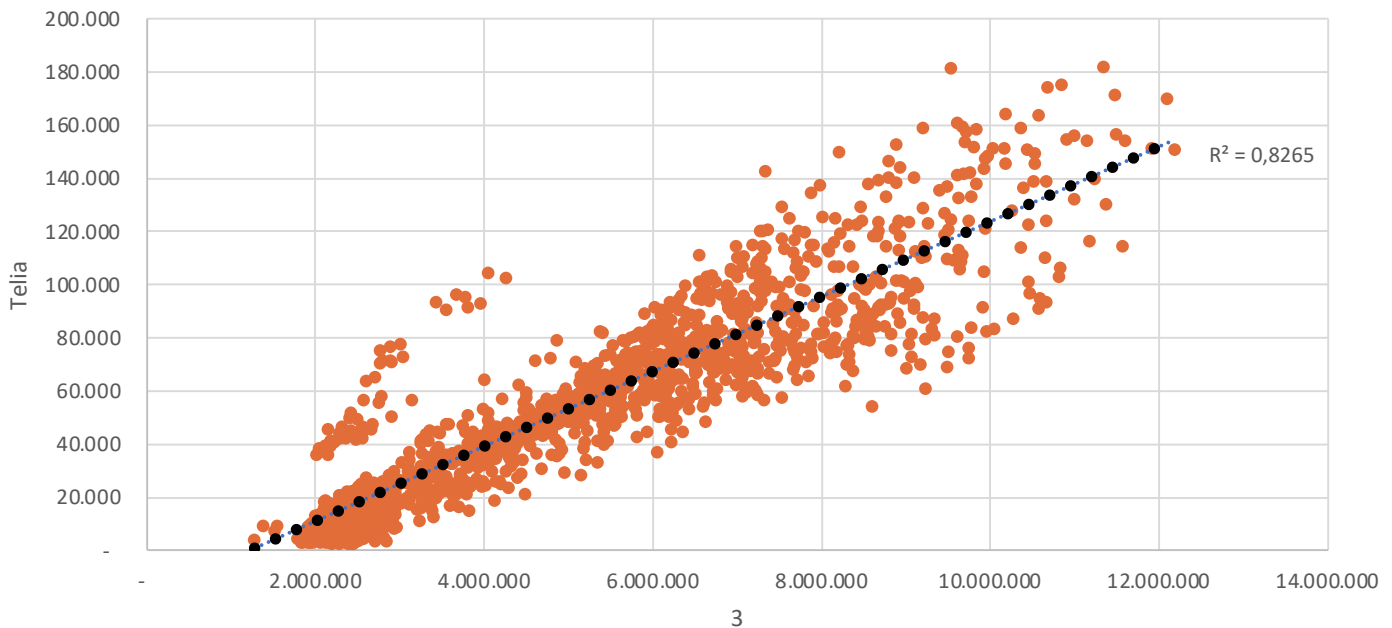
Skyscanner data Data fra flysøgninger i hele verden til København på dagsniveau. Data kan ikke brydes ned på bevægelsesmønstre i byen, men sammenligning er lavet med henblik på at sammenligne det totale antal af besøgende i byen på et givent tidspunkt og se, om der er en korrelation mellem disse tre datasæt

Benchmarking Alliance Data på overnatninger i hoteller i byen på dagsniveau, både opdateret dagligt og udgivet på månedlig basis uden de 2 måneders forsinkelse som Danmarks Statistik har på deres datasæt

Data fra 3's og Telias datasæt er desuden sammenlignet med hinanden, med henblik på at måle graden af korrelation mellem disse to datasæt, fordi der forventes at være en korrelation mellem disse to datasæt, når begge datasæt forsøger at måle på de samme KPI'er.

Figur 1: Sammenligning mellem observationer i 3s datasæt og Telias datasæt

Korrelation mellem Telias og 3s datasæt



KORRELATIONSANALYSE

Ovenstående figur er et punktdiagram, hvor hvert punkt repræsenterer en given time på en given dag i en given måned for henholdsvis Telias og 3's datapunkter. Den sorte linje er en tendenslinje, som viser hvor punkterne burde ligge, hvis de to datasæt var fuldstændigt korreleret. R2 værdien er korrelationskoefficienten, som kan variere imellem -1 og 1. Ved en værdi på 1 betyder det at de to datasæt er fuldstændig korreleret, altså hvis værdierne i det ene datasæt stiger med 1, så ville værdierne i det andet datasæt også stige med 1. En korrelationskoefficient på -1 ville vise det modsatte, altså ved en stigning på 1 i det ene datasæt ville man se et fald i det andet datasæt på 1.

En korrelationskoefficient på 0,83 viser altså, at der er god korrelation mellem de to datasæt, hvilket kan indikere, at de måler på det sammen med nogenlunde ens effektivitet. For 3's data i ovenstående graf er brugt det totale antal besøgende, altså hverken brudt ned på nationalitet eller besøgstype. Sammenligner med Telias data, som ikke kan brydes ned på besøgstype, med 3's data brudt ned på besøgstyper, får man korrelationskoefficienter som svinger mellem 0,43 og 0,83 med en enkelte undtagelse, hvilket er datasættet for lokale i 3's datasæt, som har en korrelationskoefficient på 0,09. Når data brydes ned på besøgstyper, opnås de bedste korrelationskoefficienter ved at sammenligne med datasættet for pendlere og korttidsbesøgende.

Delkonklusion 1 Umiddelbart god korrelation mellem 3's og Telias datasæt, hvilket tyder på, at de måler på det samme, og at deres metoder er forholdsvis sammenlignelige.

STORE VARIATIONER I ANTAL OBSERVATIONER

Som grafens X- og Y-akse viser, er der dog signifikant forskel på størrelsesforholdet mellem værdierne i henholdsvis Telias og 3's datasæt. Hvor Telias beregning af besøgende giver værdier mellem 2.500 og 182.000, så giver 3's datasæt værdier mellem 1.270.000 og 12.200.000, når man kigger på totalen på tværs af besøgstyper i 3's datasæt. Ser man på det totale antal besøgende i begge selskabers datasæt, er værdierne i 3's datasæt således – i gennemsnit – 143,5 gange højere end de tilsvarende værdier i Telias data på en given time / dag / måned.

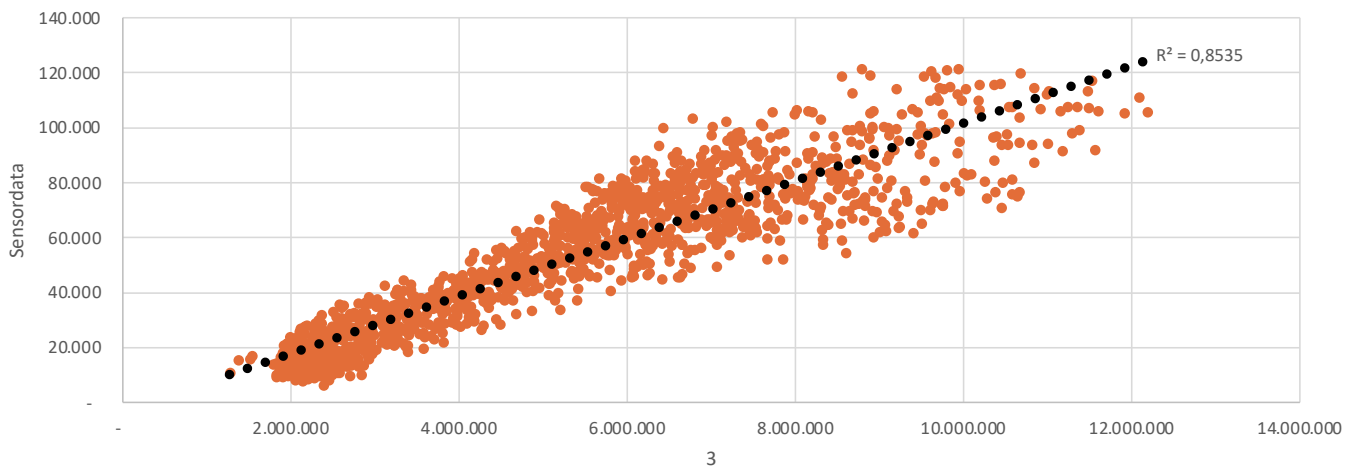
Delkonklusion 2 Meget stor forskel i størrelsesforholdet på værdierne i de to datasæt, hvilket gør det svært at bruge de absolutte værdier som grundlag for konklusioner eller indikationer. Data bør derfor, som udgangspunkt, bruges til at ses på besøgs mønstre frem for absolutte besøgstal for destinationen.

Sammenlignes datasættene fra Telia og 3 med data fra WIFI-sensorer i reklameskilte og ved andre udvalgte områder i København, ses der en flot korrelation mellem både Telias og 3's data og den data, som vi har adgang til gennem sensorerne. Dette giver mening, da alle tre datasæt forsøger at måle det samme, nemlig bevægelse i København by. Det giver således en indikation på, at datasættene kan bruges individuelt eller i kombination med hinanden til at indsamle viden om, hvordan besøgende i byen bevæger sig rundt i byen og dels, i hvilken grad de individuelle områder i byen kan være påvirket af gæstetrykket. Den viden kan, i en kontekst af turismeudvikling, bruges til at skabe strategier til at sprede turister i byen og dermed sikre at presset på et givent område ikke bliver for stort. I gennemsnit ligger Telias værdier 1,02 gange over de tal målt af sensorerne i byen og dermed svarer de to datasæt godt til hinanden størrelsesmæssigt. Derimod ligger værdierne i 3's datasæt i gennemsnit 111,84 gange over værdierne i datasættet med sensor-data og der er altså meget stor forskel på størrelsesforholdet mellem 3's data og data for sensorerne i byen, ligesom at der er meget stor forskel på 3's data og Telias data i forhold til størrelsesforholdet mellem værdierne i de to datasæt.

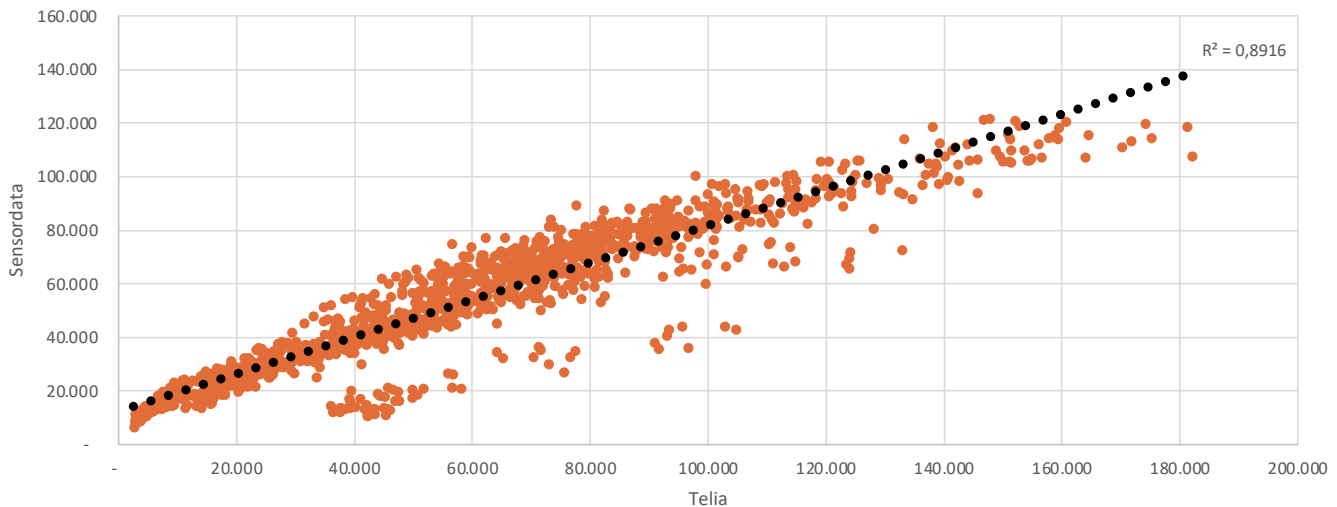
Nedenstående figurer viser sammenhængen mellem henholdsvis 3's data & sensordata og Telias data & sensordata:

Figur 2: Korrelation mellem 3 og sensordata

Korrelation mellem sensor data og 3s data

**Figur 3: Korrelation mellem Telia og sensordata**

Sensor data sammenlignet med Telia data



Delkonklusion 3 God korrelation mellem data fra WIFI-sensorer i reklamestandere og ved udvalgte punkter i København og teledata tyder på at alle tre datasæt kan bruges som en indikation på besøgendes bevægelse rundt i byen såvel som presset på individuelle områder i byen baseret på volumen af enheder i et givent område.

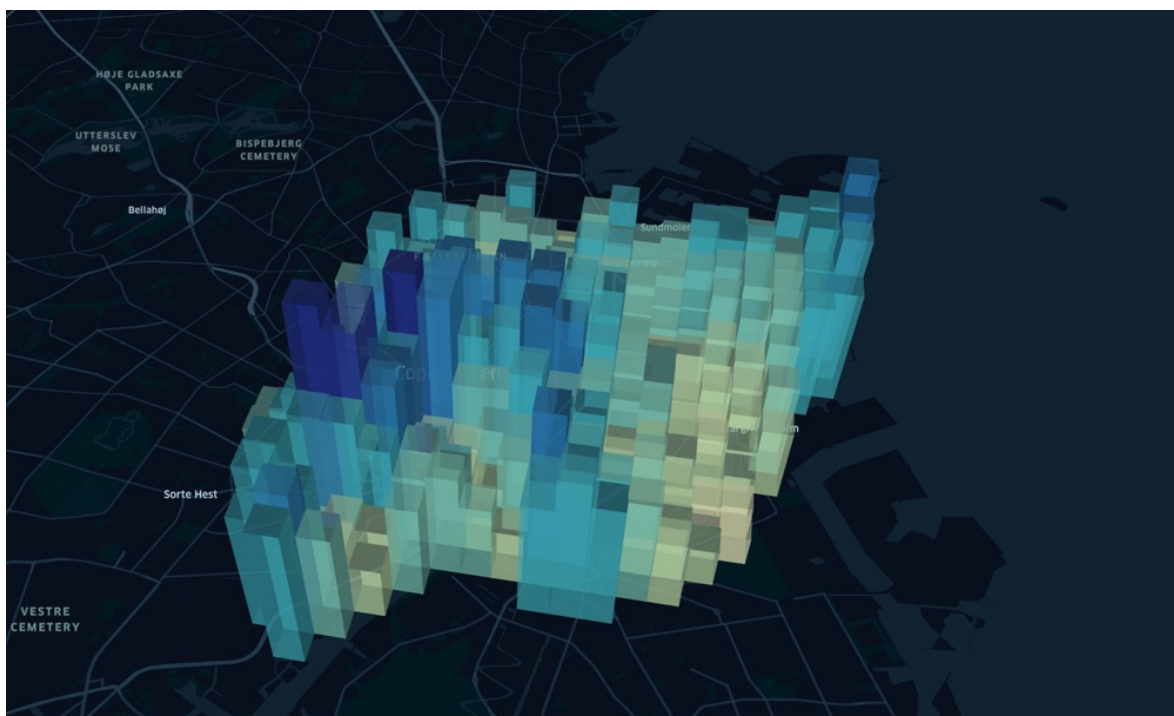
Telias og 3's data er yderligere sammenlignet med henholdsvis data fra Skyscanner, som måler på antallet af daglige søgninger på flyafgange til København fra omverdenen, og data fra Benchmarking Alliance, som måler antallet af daglige overnatninger på hoteller i København by. Det var dog ikke muligt at finde en brugbar korrelation mellem disse datasæt, idet korrelationerne mellem disse to datasæt og de to tele-datasæt lå mellem 0 og 41.

Dog skal det nævnes at disse data ikke forsøger at måle på de samme parametre. Teledata og sensordata forsøger at måle antallet af enheder (simkort, wifi-devices) i forskellige områder af byen på en given time / dag / måned, imens at Skyscanner måler på antallet af søgninger på flyrejser, som ikke alle kan antages at være konverteret til faktiske ankomster i byen, og Benchmarking Alliance måler på antallet af overnatninger på specifikke overnatningssteder i byen, men ikke på den videre brug af byens områder og gæsters bevægelse heri.

Delkonklusion 4 Der er umiddelbart ingen brugbare korrelationer imellem teledata og data fra Benchmarking Alliance eller SkyScanner, men disse datasæt forsøger heller ikke at måle på de samme KPI'er. Dog er det interessant at se korrelationen mellem disse og der kan stadig være et potentiale i at kombinere disse i en forecasting-model for besøgende i et givent område. Dog bør dette undersøges nærmere for at vurdere om de kan kombineres til dette formål.

I og med at både Telia og 3 måler på mindre områder i byen end de tilgængelige datasæt på overnatninger fra både Danmarks Statistik og Benchmarking Alliance, kan disse bruges til at give en god indikation af niveauet af besøgende i København by, såvel som deres bevægelse rundt i byen hen over en given tidsperiode. Med henblik på at give en indikation af volumen af besøgende i et givent område, kan man med fordel gøre brug af heatmaps, enten i 2D eller 3D format:





Delkonklusion 5 Teledata, givet korrelationen mellem Telias og 3s datasæt samt korrelationen med data fra WIFI-sensorer, kan umiddelbart bruges til at give en god indikation af volumen af besøgende i et givent område af byen på et givent tidspunkt. Dog er det vigtigt at huske, at data bør analyseres som relative andele og ikke i absolutte tal. Det er ikke pt. muligt at validere de absolutte tal i datasættene fra Telia og 3, man kan derimod bruge andelen af den samlede volumen i et felt, på et givent tidspunkt, til at få en indikation af, belastningen i et givent område på et givent tidspunkt. Det ville skabe potentiale for at sprede de besøgende ud og dermed mindske gæstetrykket i særlige områder.

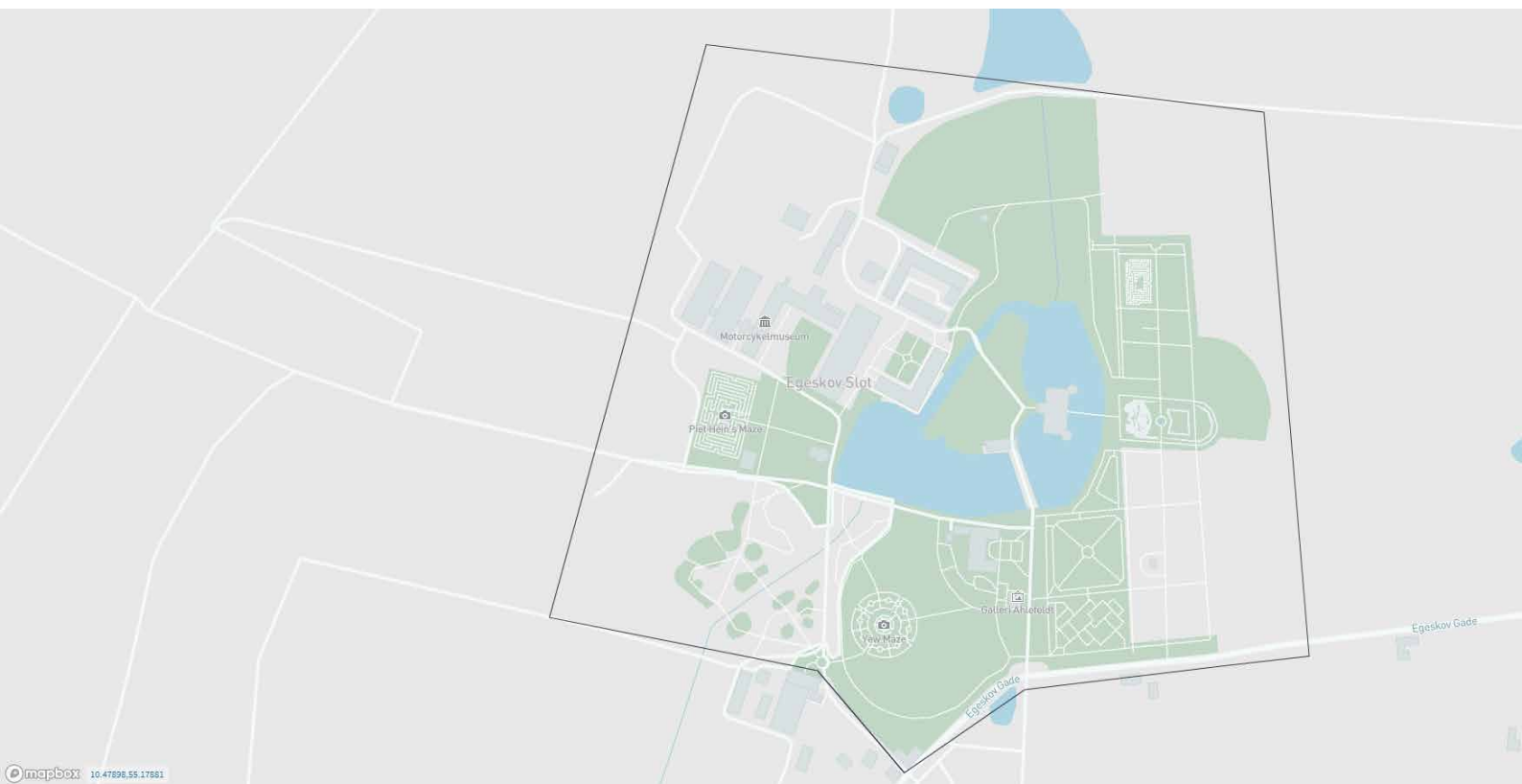
3.1.2 SAMLET KONKLUSION PÅ CASEN

De gode korrelationer mellem teledata og WIFI-sensor data tyder på, at teledata har et stort potentiale for at måle relative volumener af besøgende på tværs af områder i Danmark. Da de absolutte tal ikke kan valideres, bør der dog ses bort fra absolutte tal, så der ikke drages en konklusion om at "X besøgende var i Y område på tidspunkt Z", men at konklusioner på baggrund af teledata derimod formuleres som at data indikere at områderne 1, 2, 3 har en relativt større andel af den samlede volumen af besøgende i området på tidspunkt X end områderne 4, 5, 6.

I og med at teledata kan brydes ned på meget mindre områder, end hvad er tilfældet for data på overnatninger, har teledata ligeledes et relativt stort potentiale i forhold til at kunne visualisere små, individuelle områders andel af den samlede volumen af besøgende i byen på en let forståelig måde.

3.2 EGESKOV SLOT

Egeskov slot er beliggende på det sydlige Fyn, og er en af Danmarks bedst bevarede renæssanceborge. Slottet rummer en lang række udstillinger om dets mere end 450 års historie, og er omkranset af en stor have, med en 300 år gammel labyrint som et af dens vartegn. Slottet og haven tiltrækker årligt mere end 280.000 betalende gæster, og lægger desuden græs til mange andre kulturelle arrangementer samt festivalen Heartland.



Formålet med casen var at undersøge, hvordan teledata kunne bruges til at afdække besøgstal på et lille afgrænset område, som en attraktion, strandområde, park eller lignende. De primære KPI'er og spørgsmål som ønskes belyst var

- Niveauet af besøgende i løbet af dagen, ugen og måneden.
- Hvor kommer gæsterne fra, kommuner i Danmark og nationalitet.
- Niveauet af mennesker hen over dagen, time for time.
- Opholdslængden på attraktionen.

3.2.1 VARIABLE

For at belyse de 4 KPI'er er der brugt 2 dataset fra hver leverandør. Et datasæt som dækker de daglige antal besøgende, og et der dækker besøg pr. time.

Til at validere data er de faktiske besøgstal for Egeskov slot i perioden brugt. Data fra Egeskov slot indeholder både betalende gæster, årskortholdere samt gratister, men ikke medarbejdere, leverandører lokale, mfl. som også figurerer i teledata. Det anses stadig som en meget valid datakilde til validering af datakvalitet og validitet.



Datasæt 1: Egeskov - Besøg pr. Dag

	3	Telia	Egeskov
Periode for data	01-06 til 31-08-2022	01-07 til 31-08-2022	01-06 til 31-08-22
Dato	x	x	x
Homezone (hjemme-kommune)	x	x	
Antal besøg	x	x	x
Antal unikke besøg	x	x	
Gæstetype	x		
Opholdslængde	x		

Datasæt 2: Egeskov - Besøg pr. time

	3	Telia
Periode for data	01-06 til 31-08-2022	01-07 til 31-08-2022
Dato	x	
Time	x	x
Ugedag		x
Homezone		
Antal besøg	x	x
Antal unikke besøg	x	x
Gæstetype	x	
Opholdslængde	x	

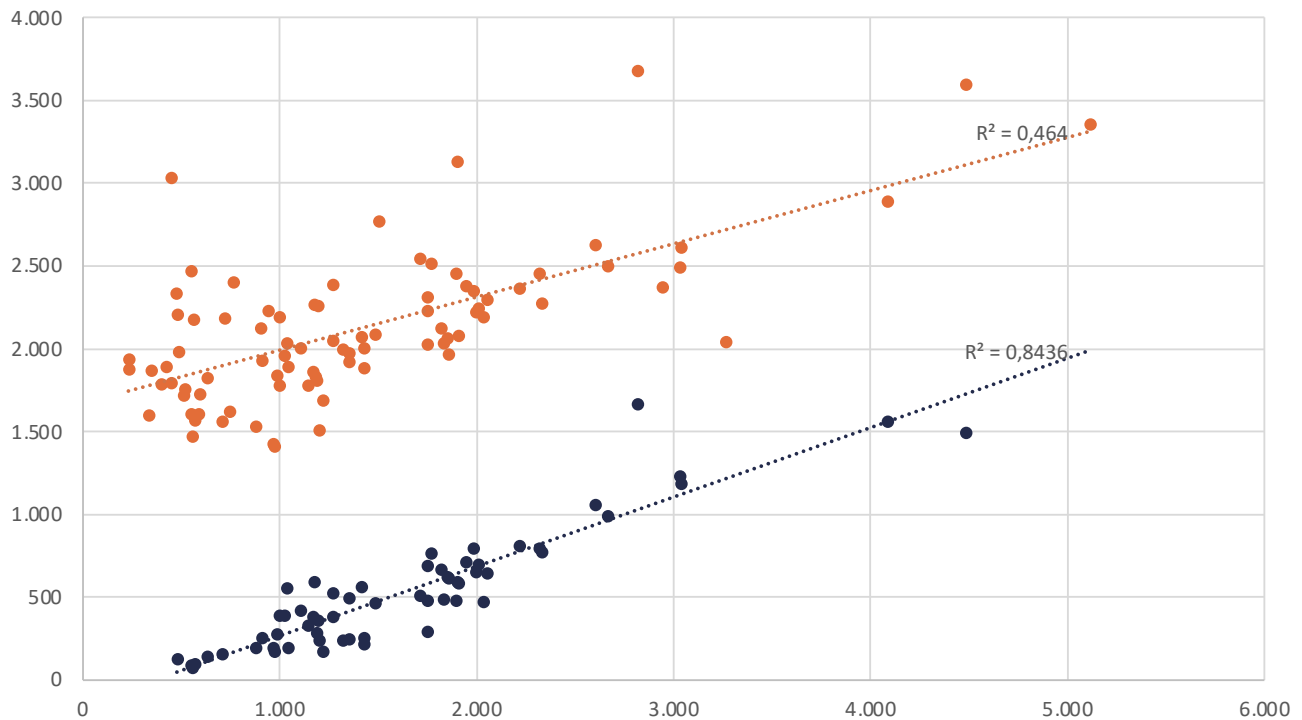


Som i projektets øvrige cases, er den store forskel på data fra 3 og Telia, at 3 inddeler deres besøg i 6 gæstetyper, samt medtager internationale gæster, mens Telia kun medtager danskere og uden at skelne til besøgstype. Begge extrapolerer deres data så det er repræsentativt i forhold til den danske befolkning.

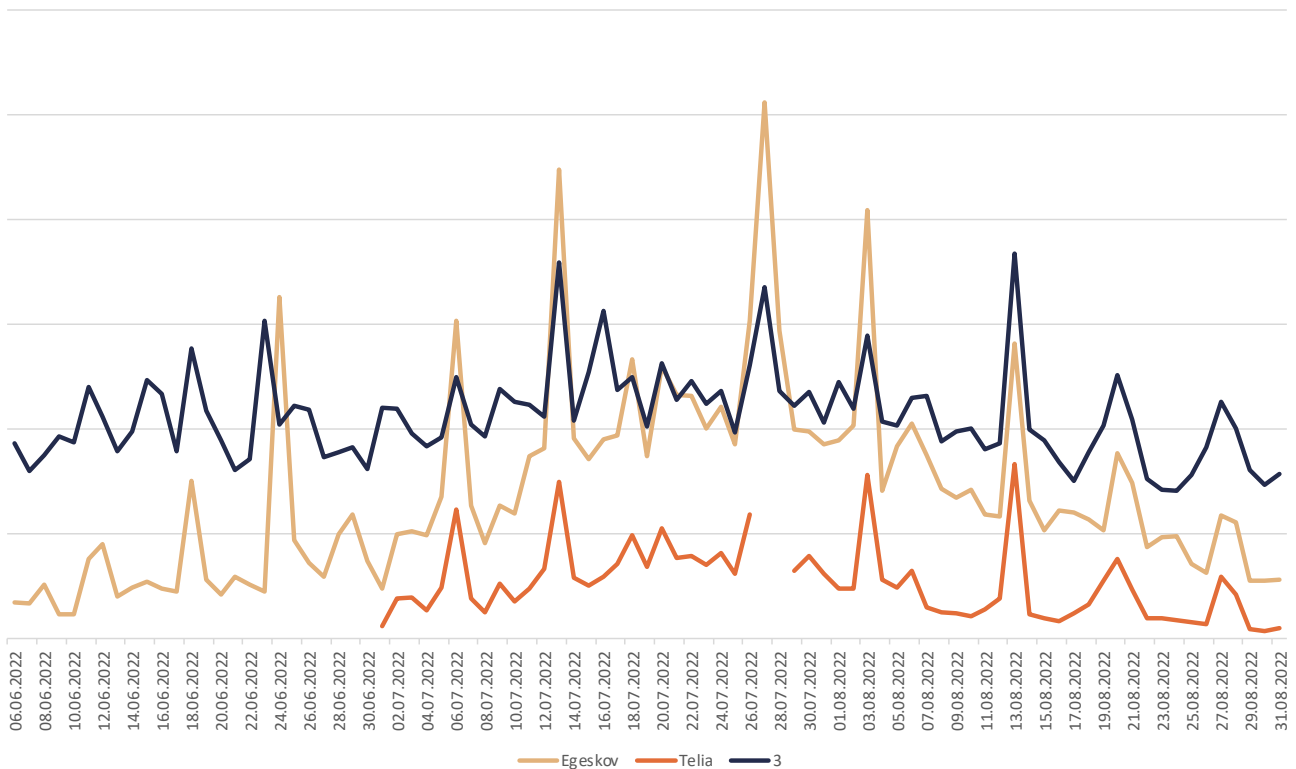
Kigger vi nærmere på de forskellige gæstetyper for 3, ses et meget stort antal af transit, med mere end 28.000 i gennemsnit pr. dag. Samtidig er niveauet af beboere på lidt over 3.300 pr. dag hvilket tyder på, at det er mere end de omkringliggende huse som medtages. Dette understøttes yderligere af at der er registreret beboere fra Svendborg, hvor kommunegrænsen ligger ca. 3 km. fra Egeskov. Tegnes en 3 km. radius fra Egeskov, fanges byen Kværndrup, som har ca. 1.700 indbyggere, og både hovedvejen, og Svendborgmotorvejen går igennem dette område, hvilket muligvis kan forklare det store antal transit gæster.

3.2.2 VALIDERING AF KPI'ER

Til at sammenligne data bruges alene kategorien med endagsbesøgende fra 3, mens det fulde datasæt bruges for Telia, da de ikke differentiere imellem besøgstype. Som valideringskilde bruges besøgende på Egeskov slot, hvor der her både medtages almindelige betalende gæster, gratister og årskortholdere.

Figur 4: Korrelation mellem teledata og valideringskilden**NIVEAUET AF BESØGENDE I PR. DAG, UGE OG MÅNED.**

Som det kan ses af figur 5, er der ingen af de to målinger som rammer præcist. De har dog begge to en relativ god korrelation, og følger trenden i besøgende hen over uge og måned. Telias tal ligger generelt lavere end valideringskilden, med en gennemsnitlig daglig præcision på 31%, mens 3 generelt ligger over valideringskilden med en gennemsnitlig daglig præcision på 217%. Altså, antallet af besøgende beregnet af Telia ligger, i gennemsnit, på en tredjedel af de officielle besøgstal fra Egeskov, imens at 3, i gennemsnit, beregner dobbelt så mange besøgende som Egeskovs eget data viser.

Figur 5: Antal besøgende pr. dag (uden Y-akse, efter ønske fra Egeskov)

Figur 4 og 5 viser tydeligt den store variation, der er i målingerne fra 3, der på dagsniveau svinger i præcision fra 63% til 827%, hvorimod Telias data er mere konstant med meget lavere variation. På uge niveau ligger Telia med en præcision på mellem 29%-37% hvorimod 3 svinger fra 90%-406%. På trods af den store volatilitet i målingerne rammer 3's estimeringer dog væsentligt mere præcis i ugerne 28-31, hvor gæstemængden ved Egeskov generelt er højere.

En mulig årsagsforklaring til den konstante underestimering fra Telia, kan skyldes de "manglende" internationale gæster i data. Da det ikke har været muligt at adskille danske og internationale besøgende i Egeskovs data, udgør valideringskilden derfor alene den totale gæstemængde, og vil derfor naturligt ligge højere end Telias estimering af gæster i området, som kun indeholder danske besøgende.

NIVEAUET AF BESØGENDE I PR. DAG, UGE OG MÅNED.

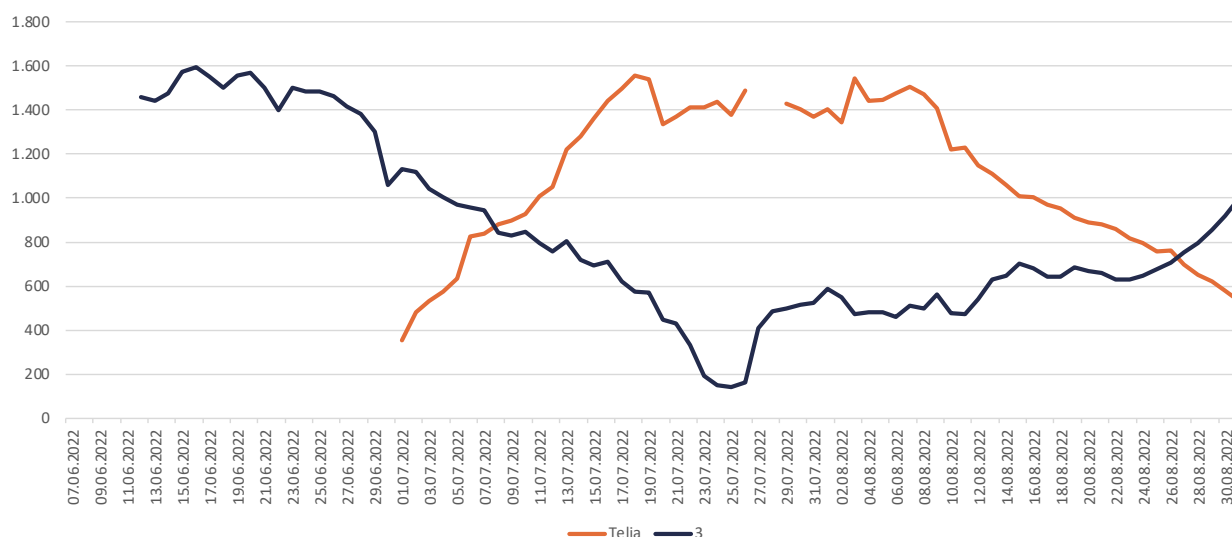
Præcision	Telia	3
Uge-23		406%
Uge-24		346%
Uge-25		213%
Uge-26		221%
Uge-27	35%	149%
Uge-28	32%	117%
Uge-29	37%	104%
Uge-30		90%
Uge-31	30%	108%
Uge-32	31%	142%
Uge-33	31%	152%
Uge-34	29%	186%

	Telia	3
Korrelationskoefficient	0.9185	0.6812
Forklaringsgrad (R2)	0.8436	0.4640
Gennemsnitlig absolut afstand	1093.2	861.2989
Præcision	31%	217%
Maks.	59%	827%
Min.	13%	63%

Kigger man på afstanden mellem data og valideringskilden på et 7 dages gennemsnit, ses det, at 3 reducerer fejlmængden jo flere gæster, der besøger området. I start juni er 3 i gennemsnit 1.400 gæster fra valideringskilden, hvilket reduceres til under 200 i slutningen af juli, hvorefter det stiger igen. Telia øger blot fejlmængden i takt med antallet af gæster stiger, hvilket muligvis kan tilskrives, at Telias data ikke registrerer internationale gæster, og at det netop er denne andel, der stiger i højsæsonen, hvorfor fejlmarginen øges. Udsvingene i 3's data kan ikke forklares ud fra informationerne leveret til denne rapport, men kan antages at skyldes, at usikkerheden i 3's estimeringen af antal gæster falder, når antallet af gæster i et område stiger.

Figur 6: Fejlmargin fra besøgstal i valideringskilden (7 dages glidende gennemsnit)

Afstand til valideringskilde, 7 dages glidende gennemsnit



Hvor kommer gæsterne fra, kommuner i Danmark og nationalitet.

Både 3 og Telias data angiver, hvilken kommuner de danske gæster kommer fra, og i nedenstående analyse vises de største regioner samt aggregeret data på regionsniveau. 3's data vises kun for kategorien endagsbesøgende.

Fordeling af besøg	Telia	3
Syddanmark	82%	84%
Midtjylland	6%	7%
Sjælland	3%	5%
Hovedstaden	6%	3%
Nordjylland	2%	1%

Kigger man på antallet af gæster fordelt på regioner har de begge en meget ens fordeling, hvor den største forskel på 3 procentpoint er for Region Hovedstaden. Forskellene øges dog, når data brydes yderligere ned på måneds, uge og dagsniveau, og antallet af observationer fra de enkelte kommuner bliver mindre.

Inddeler man på kommuneniveau i stedet for, ses en større forskel i fordelingen på især de fynske kommuner og særligt for Faaborg-Midtfyn som Egeskov ligger i, og Svendborg kommune hvor kommunegrænsen kun ligger 3 km. væk fra slottet. 3 angiver at 41% kommer fra Faaborg-Midtfyn, og 17% fra Svendborg, mens det for Telia er hhv. 19% og 18%.

Datasættet fra 3 indeholder også internationale gæster. Til validering af dette er brugt overnatningstallene fra Danmarks statistik (DST), da det ikke var muligt at få besøgstal fra Egeskov fordelt på nationalitet. Datasættet fra 3 viser en fordeling mellem danske og internationale på 92% / 8%, hvor splittet fra DST er 65% / 35%. Ser man på fordelingen på nationaliteter, så udgør Tyskerne langt størstedelen af de internationale besøg, med 64%, hvilket matcher overnatningstallene ret godt da tyskerne udgør 60% af de internationale overnatninger. Holland er for begge to det næststørste marked med en andel på 6% for 3, og 14% for DST. Herefter udgør Norge 7% af overnatninger men kun 1% af besøgende fra 3. Det er svært at konkludere noget direkte i forhold til fordelingen. Overordnet set er andelen af internationale i 3's data for lavt, og den tyske andel af gæster er derfor svært at konkludere på.

Nationalitet	3	DST-Overnatninger DestinationFyn
Danmark	91.88%	65.4%
Internationale	8.12%	34.6%
Tyskland	64.3%	60.1%
Holland	6.0%	14.3%
Belgien	5.9%	1.7%
Sverige	5.3%	4.0%

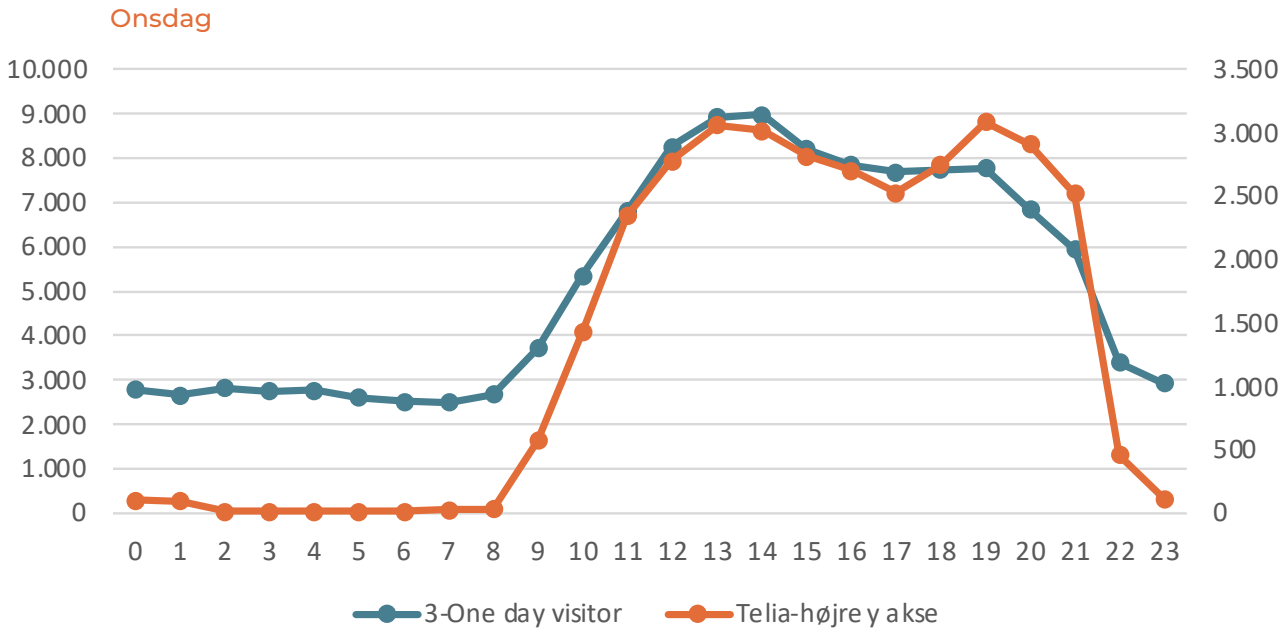
NIVEAUET AF GÆSTER FORDELT PÅ DAGEN, TIME FOR TIME.

Den tredje KPI "hvordan endagsbesøgende fordeler sig hen over dagen" ses akkumuleret for alle ugedage i figur 7. Nævneværdigt er det store antal registrerede endagsbesøgende udenfor åbningstid og særligt om natten i 3's data. I figur 8 ses det desuden, at der er forskelle i mønstret på kategorierne som 3 anvender. Det kunne tyde på, at gruppen med endagsbesøgene, der "besøger" Egeskov i tidsrummet før kl. 09.00 og efter 20.00, er fejkategoriseret og ikke burde indgå. Figur 8 viser, med andre ord, hvilke af 3's kategorier af besøgende som er ansvarlige for at 3's data viser et antal af besøgende på Egeskov slot udenfor slottets åbningstid. Der kan mere eller mindre ses bort fra tallene for beboere, fordi 3's datasæt inkludere et geografisk område med beboelse og disse tal er forholdsvis stabile på tværs af tidspunkter, men de andre kategorier bidrager i større eller mindre grad til antallet af besøg udenfor slottets åbningstid, herunder en relativt stor andel af kategorien endagsbesøgende.



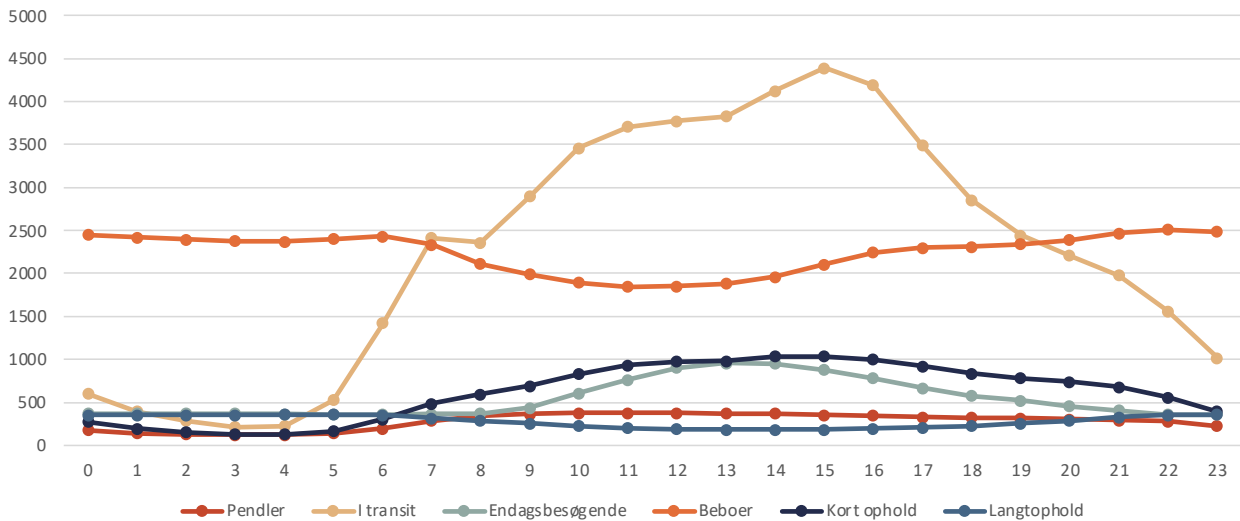
Foto: Skagen, Mette Johnsen

Figur 7: Akkumulering af observationer på tværs af alle dage, fordelt pr. time

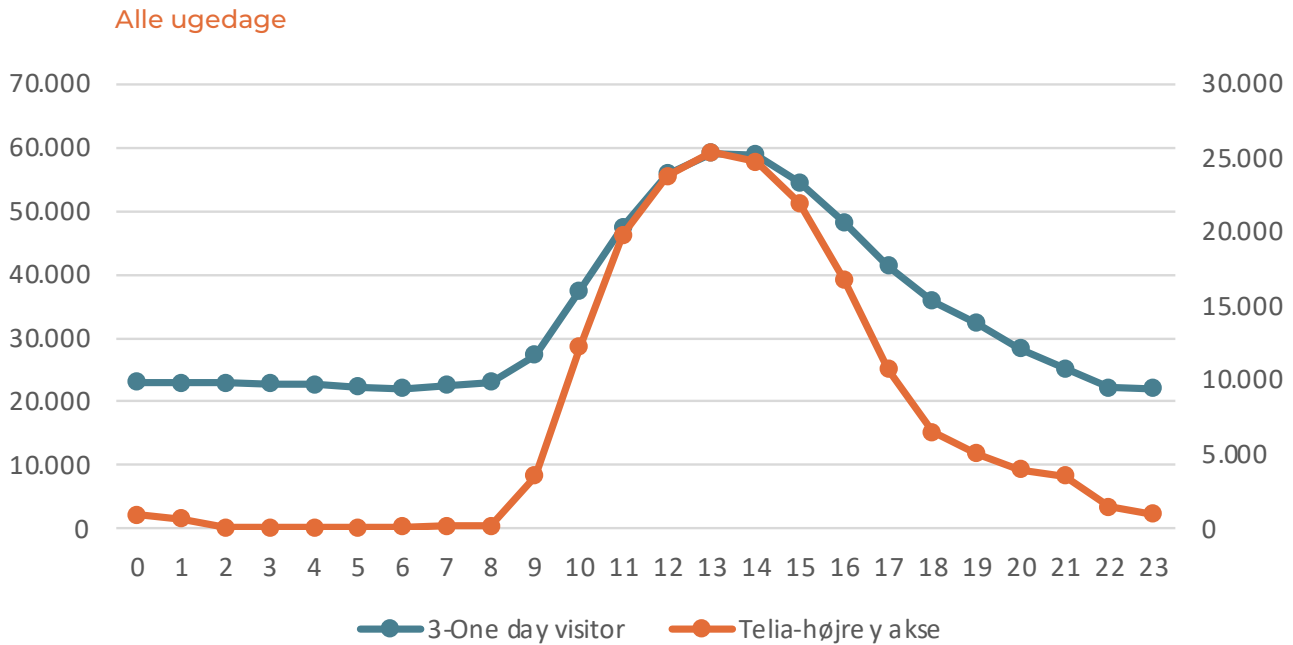


Figur 8: Gennemsnitligt antal observationer pr. gæstetype, fordelt pr. time

3 Gns. observationer pr. timer pr. kategori



Samme mønster ses hen over alle de individuelle dage, blot forskudt om onsdagen i figur 8, hvor Egeskov havde udvidet åbningstider til kl. 22.00. Dette kan tydeligt aflæses i graferne nedenfor. Begge teledataudbydere fanger dog fint den øgede aktivitet mellem kl. 19.00-22.00 på onsdage.

Figur 9: Akkumulering af observationer på tværs af alle onsdage, fordelt pr. time**OPHOLDSLÆNGDEN PÅ ATTRAKTIONEN**

Opholdslængden er kun en del af datasættet for 3, og er derfor svært at validere. Kigger man på opholdslængden for de forskellige kategorier, ser det umiddelbart rigtigt ud, og passer til hvad man kunne forvente af de forskellige typer gæster.

	Opholdslængde, Timer
Beboer	4.18
Pendler	1.76
I transit	0.06
Endagsbesøgende	3.88
Langt ophold	58.43
Kort ophold	0.60

Kigger man specifikt på kategorien endagsbesøgende, så har de en gennemsnitlig opholdslængde på 3,9 timer, som må anses at være meget validt i forhold til, hvad man kunne forvente. Der næsten ingen forskel på opholdslængden, når man ser på tværs af ugedage, som varierer mellem 3,7 timer om mandagen til 4,1 time om søndagen. Kigger man hen over dagen er det også meget lille forskel på opholdslængde, dog er de, der ankommer mellem kl. 09-11, der i gennemsnit mere end 4 timer.

3.2.3 KONKLUSION PÅ CASEN

Samlet set så er korrelationen mellem data og valideringskilden ganske god, men niveauet af observationer er meget svingende og upræcist. Man kan derfor kun sige noget om trenden og det relative antal gæster – det absolutte antal er for usikkert. Dette gælder både, når man kigger på timeniveau og på dags- og ugeniveau.

På trods af et defineret afgrænset POI tyder det på, at det har været svært at isolere det lille område, og særligt data fra 3, ser ud til at have medtaget observationer i et område, der er markant større, hvilket kommer til udtryk, når de inddeler observationerne i gæstetyper. Derudover er der et relativt højt antal observationer af endagsbesøgende udenfor åbningstiderne, hvilket kunne tyde på, at de er placeret i den forkerte gæstetype, da Egeskov ikke har overnatningsfaciliteter. Telia har derimod meget stabile tal og fordelinger over døgnets timer. De har dog en konstant underestimering af antal gæster, der formentlig er forårsaget af den "manglende" registrering og ekstrapolering af internationale.

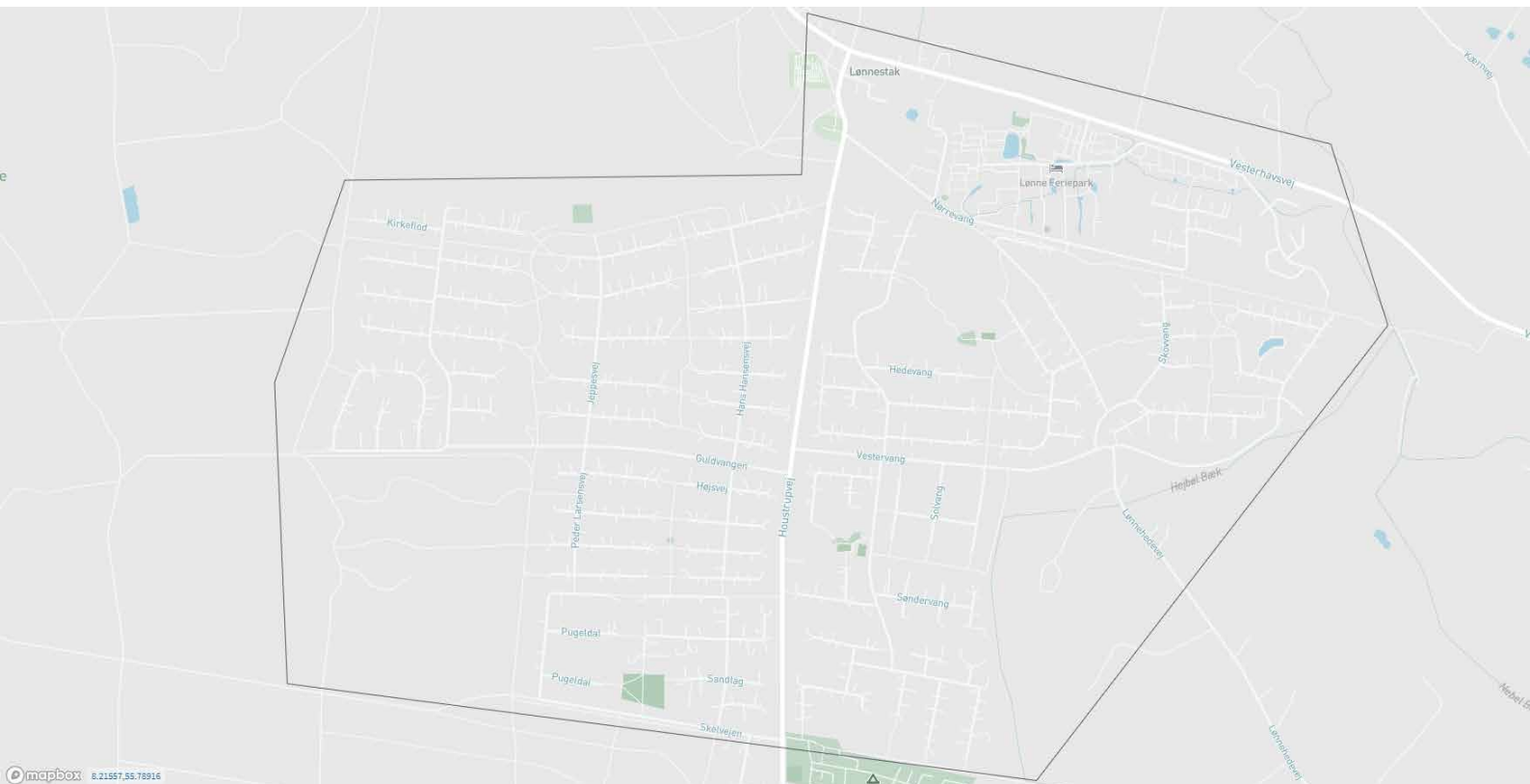
Et andet generelt fund er, at jo større område/gruppering der sammenlignes i data, des større sammenfald er der mellem de to teledataleverandørers data. Kigger man fx på regionsfordelingen, er der meget lille afvigelse mellem de to. Hvorimod der i opdelingen af, hvilken kommune gæsten kom fra (hjemmezone) – og særligt kommunerne lige omkring Egeskov; Faaborg-Midtfyn og Svendborg – er der stor variation, hvilket kan skyldes 3's kategorisering af gæstetyper i data eller forskelle i ekstrapoleringsmetoderne.



Foto: Tisvilde Hegn, Daniel Villadsen

3.3 LANDAL SEAWEST

Landal Seawest er et feriecenter beliggende i Houstrup i Nørre Nebel kommune. Landal ligger side om side med et større sommerhusområde. Casen omhandler både Seawest samt det meste af sommerhusområdet, da Landal alene ikke kunne isoleres i teledatakilder. I området løber derudover en række gennemkørselsveje samt en mindre Campingplads syd for POI'en.



Casens formål er at validere teledata for et ferieområde præget af markante sæsonudsving, men i udgangspunktet ikke påvirket af enkeltstående events.

Følgende KPI'er er forsøgt belyst i casen:

- Antal gæster i POI'en – og hvilke kommuner kommer de fra (antal + oprindelse)

Følgende KPI'er er undersøgt, men ikke belyst i denne case grundet datakvaliteten af ovenstående KPI:

- Antal gæster i POI'en, der rejser ud af POI'en under deres besøg (mobilitet + endagsturisme i omkringliggende kommuner)
- Antal gæster, der besøger X andre kommuner under et besøg (antal besøgs-kommuner)
- Opholdslængder (mobilitet)

3.3.1 VARIABLE

Casen indeholder data fra 3 og Telia samt valideringsdata baseret på sommerhusbookinger fra Houstrup-området (leveret af NTT) og data fra Landal Seawest (forkortet SWT).

Landal – besøg pr. dag

	3	Telia	NTT	SWT
Periode	01-06-22 til 31-08-22	01-07-22 til 31-08-22	01-06-22 til 31-08-22	
Dato	X	X	X	X
Nationalitet	X		X	X
Kommune	X	X		
Type af besøgende	X			
Antal besøgende	X			
Antal unikke besøgende	X	X		
Returnerede besøgende	X	X		
Besøgende til X	X	X		
Bookinger (antal personer)			X	X

NTT har kun bookingdata for en delmængde af sommerhusene i Houstrup, og data er derfor opjusteret baseret på manuel optælling af samtlige sommerhuse i POI'en (undtagen SWT-området). Med opregningen bør NTT's data være retvisende for Houstrup-området under antagelse om, at deres data er repræsentativt. NTT og SWT's data er herefter samlet til én datakilde, og udgør valideringskilden i denne case.

Bemærk at Telias data hverken indeholder internationale gæster eller datoer for juni måned.

3.3.2 VALIDERING AF KPI'ER

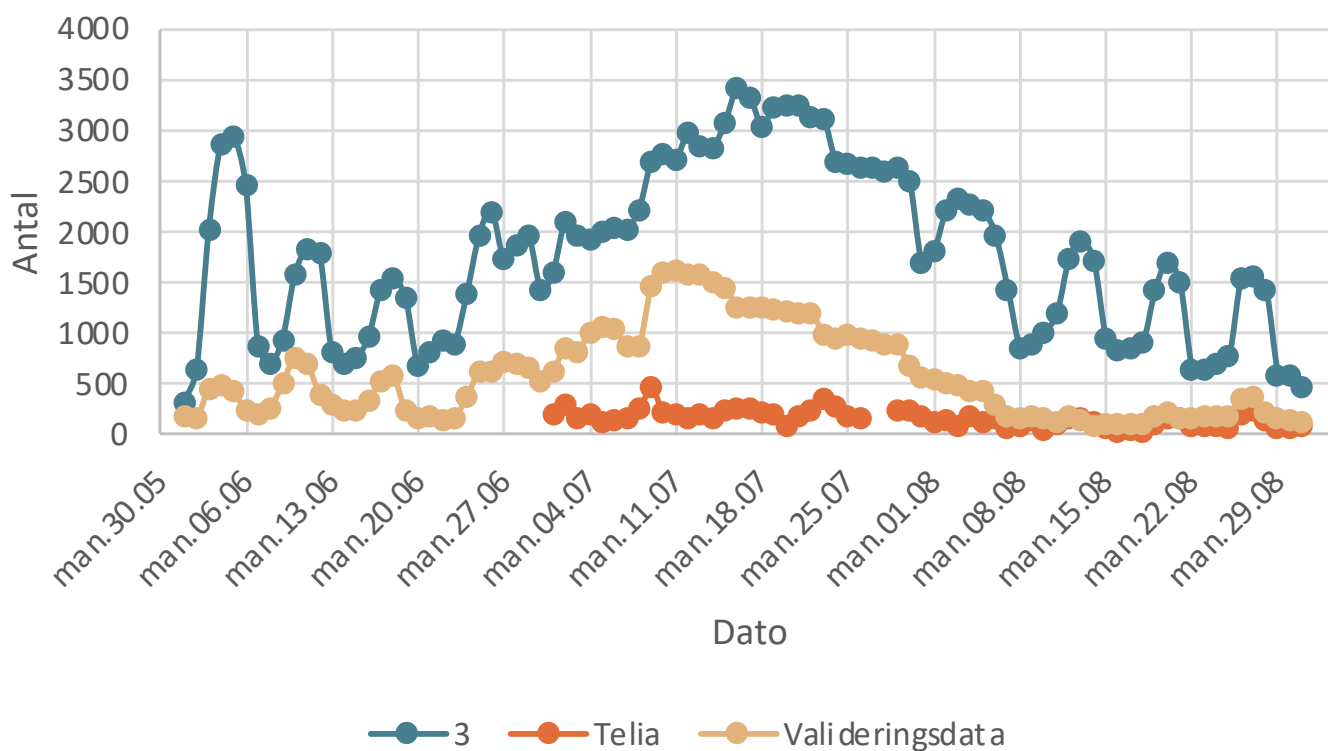
ANTAL GÆSTER

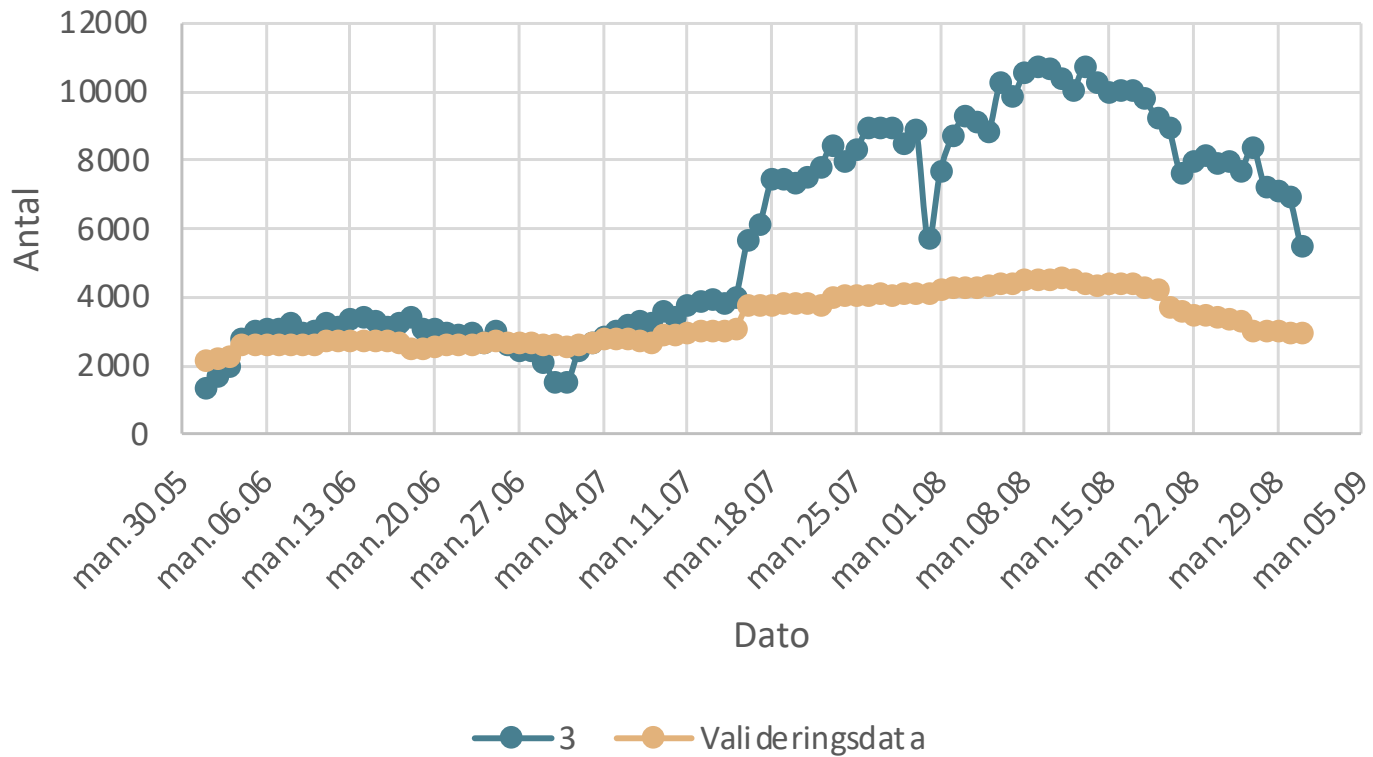
Figur 10 og 11 viser antallet af gæster/bookinger fra de forskellige datakilder fordelt på danske og internationale gæster.

I figur 10 ses der i 3's data et meget højere antal end både valideringskilden og Telia. Telia viser omvendt et lavere antal end valideringskilden i juli, imens niveauerne konvergerer for august, hvilket ikke umiddelbart kan forklares. For danske gæster er det tydeligt, at der er nogle mønstre, som går igen for både 3 og valideringskilden: Markante stigninger på ugebasis i juni og august samt en relativ stabil højsæson i juli. Det er sværere at vurdere for Telia, da tallene generelt er lave og tilsyneladende er behæftet med stor usikkerhed.

For hele perioden viser 3 i gennemsnit 1.210 flere danske gæster på dagsbasis end valideringskilden, imens Telia på dagsbasis viser 500 færre danske gæster end valideringskilden.

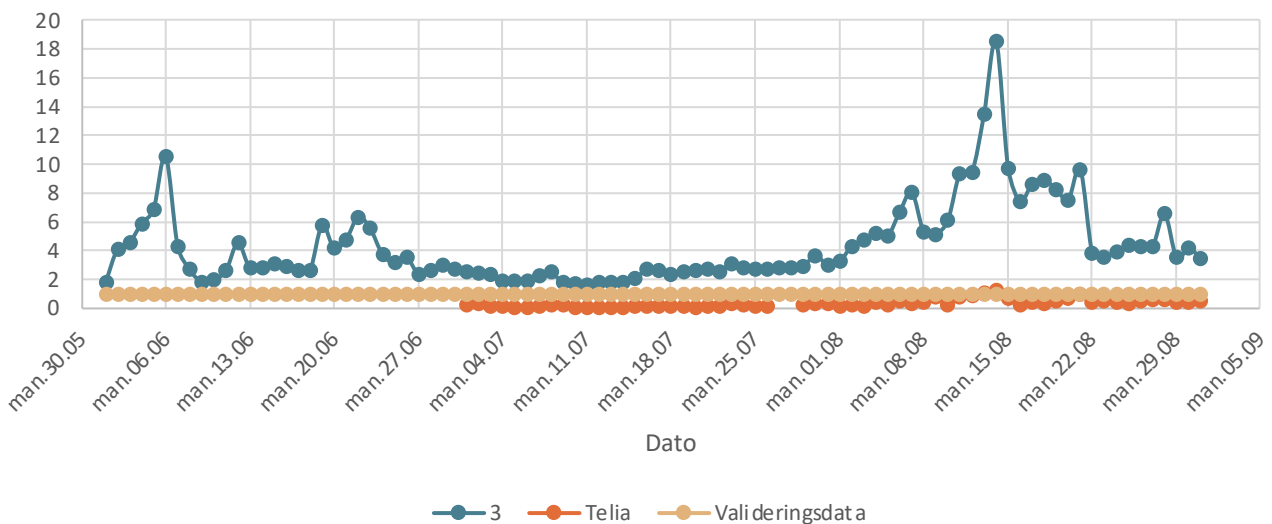
Figur 10: Danskere

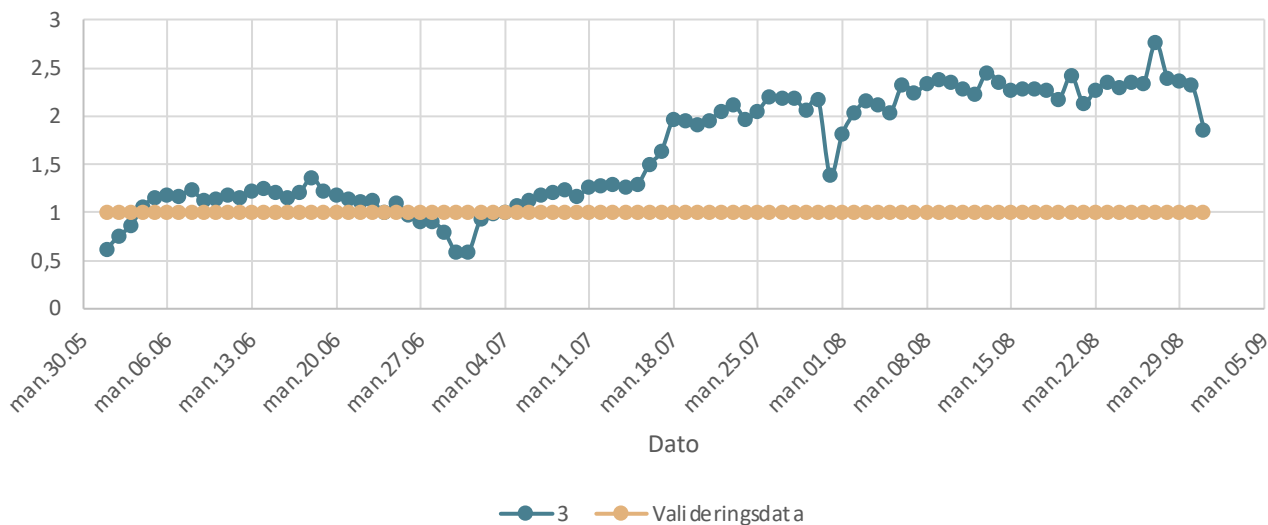


Figur 11: Internationale

I figur 11, for internationale gæster, passer antallene godt for juni og første halvdel af juli, imens differensen bliver markant for højsæsonen i anden halvdel af juli og august. For hele perioden viser 3 i gennemsnit 2.479 flere internationale gæster på dagsbasis end valideringskilden

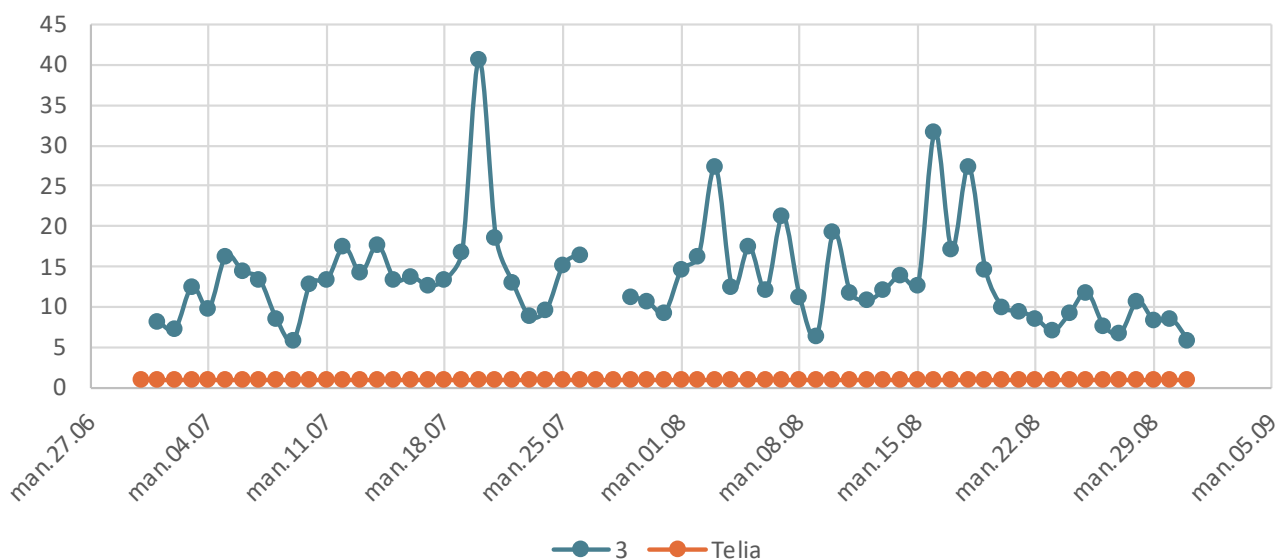
I figur 12 og 13 indekseres valideringskilden til indeks = 1, og her ses tydeligt, hvordan 3 estimerer meget høje tal. For hele perioden ligger 3's tal i gennemsnit 336% højere for danske gæster (figur 12) og 64% højere for internationale gæster (figur 13). Omvendt ligger Telias tal 61% lavere end valideringskilden for danske gæster i perioden for juli og august (figur 12).

Figur 12: Danskere, indekseret (Valideringsdata = 1)

Figur 13: Internationale, indekseret (Valideringsdata = 1)

Der er desuden lavet korrelationsanalyser mellem valideringskilden og henholdsvis 3 og Telia. Generelt er korrelationen højere for 3 end for Telia, hvilket betyder, at 3's data umiddelbart stemmer bedst overens med valideringsdata, så høje og lave værdier følger hinanden. Overordnet er der dog betydelige usikkerheder fra dag til dag for både 3 og Telia.

Sammenligner vi 3 og Telias data direkte med hinanden, bliver både afvigelserne i datatendenserne (udsving) og selve antallet underbygget: I figur 14 er Telias data indekseret til 1. 3 ligger i gennemsnit for de to måneder omkring 13,5 gange højere end Telia - når Telia estimerer 1 gæst estimerer 3 altså 13,5 gæster. Derudover er sammenhængen mellem værdierne generelt meget usikre: 3 viser mellem 5 og 40 gange så høje værdier som Telia, og niveauet er ikke konstant over perioden.

Figur 14: Teleselskaber indekseret (Telia = 1)

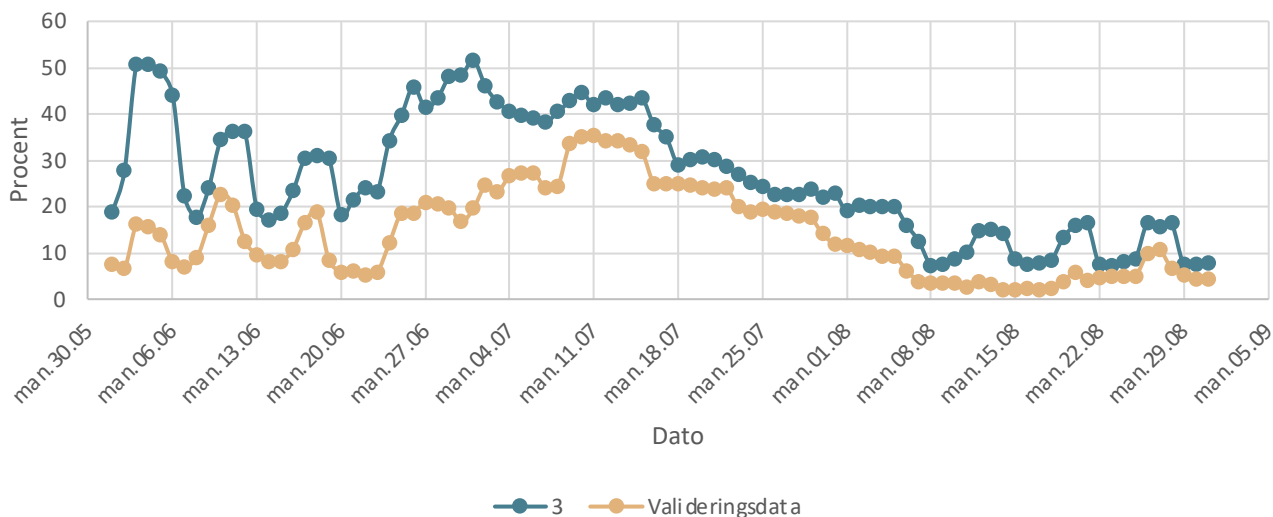
For KPI'en "antallet af gæster" er konklusionen behæftet med en del usikkerheder. 3 og valideringskilden viser en del sammenhængende mønstre hen over perioden - omend antallet i absolut forstand ligger langt fra hinanden. Dette gør sig gældende for danskere såvel som internationale gæster.

Telia viser ikke mange meningsfulde mønstre og ligger tilmed markant lavere end særligt 3, men også valideringsdata. 3 og Telia viser direkte sammenlignet også meget forskellige tal. Overordnet viser 3 den bedste sammenhæng med valideringskilden, om end tallene syner overestimerede.

En anden måde at vurdere gæstemængden på, er ved at sammenligne procentandelen af danske gæster med valideringskilden. I figur 15 sammenlignes 3's procentandel med valideringskilden, og figuren underbygger, at andelen af danske gæster ligger markant højere i 3's data for hele perioden. 3's data viser således både flere gæster samlet set end valideringskilden, men også en skæv andel af danske gæster. For hele perioden angiver 3, at 27% af gæsterne i gennemsnit er danske, imens det for valideringskilden kun er 14%, hvilket næsten svarer til det dobbelte.

Der er dog en nogenlunde konsistent sammenhæng mellem høje og lave procentandele på tværs af hele perioden, på nær en særlig høj andel i perioden 27/6 til 4/7. Selvom de eksakte gæsteantal er forskellige, viser begge kilder således klart, hvordan andelen af danskere følger et "weekendmønster" i juni og august, en mere "stabil" højsæson i hele juli, hvorefter andelen af danske gæster forventeligt falder i august.

Figur 15: Procentandel danskere pr. datakilde



HVOR KOMMER GÆSTERNE FRA?

Der er foretaget analyser for antallet af besøgende fra forskellige danske kommuner, som ikke vises her grundet for store udsving i data. Telia har data på dagsniveau, men tallene viser meget store udsving. På ugeniveau er der data for 3 og Telia, som viser et nogenlunde konsistent mønster: De fleste gæster kommer fra Region Syddanmark efterfulgt af Region Midtjylland. Analyserne er foretaget på regionsniveau, da datagrundlaget for flere kommuners vedkommende bliver for usikkert. Tallene vurderes dog stadig for usikre til at konkludere noget entydigt ud fra.

Konklusionen på KPI'en "hvor kommer gæsterne fra" er, at 3 viser markant flere danskere end valideringskilden. Udviklingen over tidsperioden er dog nogenlunde konsistent mellem kilderne på nær enkelte udsving. Det eksakte gæsteantal er altså svært at svare entydigt på, men der kan gives et kvalificeret bud på, hvornår der henholdsvis er høje/lave andele af danske og internationale gæster. Besøgsantal på kommuneniveau er generelt meget usikkert, da datagrundlaget er for småt.

3.3.3 CASE-KONKLUSION

Både 3 og Telia viser generelt tal, som ligger langt fra valideringskilden – med undtagelse af internationale gæster i juni og starten af juli for 3. 3 overestimerer generelt antallet og Telia underestimerer ganske signifikant. Direkte sammenlignet estimerer 3 langt højere tal end Telia, og forholdet er præget af store variationer. 3 og valideringskilden viser dog en pæn sammenhæng relativt set, hvor særligt weekendudsving samt høj- og lavsæson for danskere og internationale gæster er tydelige. Samlet set vurderes det, at casen ikke kan give en valid estimering af hverken det samlede antal gæster i området, nationaliteter, regioner eller kommuner. Antallet af endagsturister, der overnatter i området og besøger de omkringliggende områder som endagsturister, kan derfor heller ikke valideres.

Den samlede usikkerhed i kilden skyldes formentligt, at der er for få observationer (=mennesker) i det givne område.

3.4 RAMMSTEIN KONCERT I AARHUS

Koncerten blev afholdt i Aarhus d. 22. juni 2022 i Ceres Park Arena, og var ifølge arrangørerne udsolgt med cirka 35-40.000 gæster. Området er afgrænset til selve koncertpladsen og det nærtliggende areal rundt om stadion. I hverdagen benytter mange forskellige grupper faciliteterne, men pga. det store set-up i forbindelse med koncerten, var det meste af området lukket ned for offentligheden på selve koncertdagen.

Casens formål er at afdække anvendelsen af teledata i forbindelse med et stort event på et mindre område med deltagere fra flere forskellige geografiske områder. I casen er de følgende KPI analyseret:

- Antal gæster i POI'en – og hvilke kommuner kommer de fra (antal + oprindelse)

3.4.1 VARIABLE

Casen indeholder datakilder fra 3 og Telia. Det har desværre ikke været muligt at indhente data på koncertgæsterne via bookingbureauet, og derfor kan valideringen primært foretages ved direkte sammenligning af teleselskabernes data.

Datakilderne består af gæsteantal fordelt på kommuneniveau på timebasis for koncertdatoen d. 22. juni 2022. Kun 3's data indeholder internationale gæster.

Rammstein – besøg pr. time

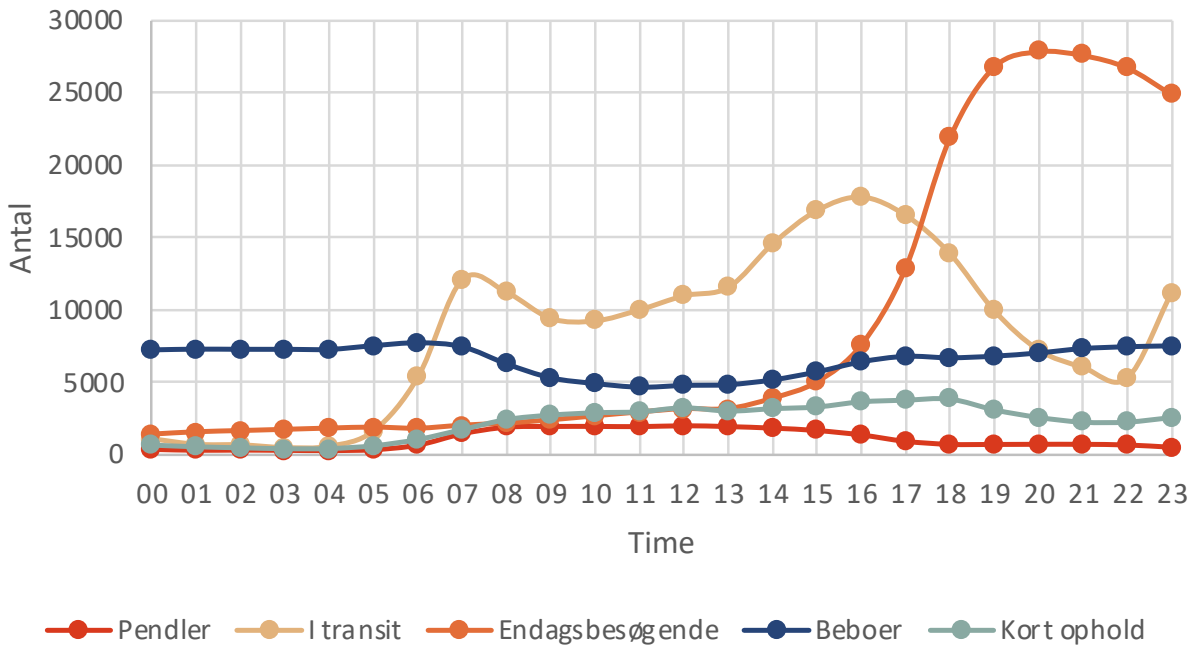
	3	Telia
Periode	22-06-2022	22-06-2022
Time	X	X
Kommune/Land	X	X
Type af besøgende	X	
Antal besøgende	X	
Antal unikke besøgende	X	X

3.4.2 VALIDERING AF KPI'ER

SAMLET ANTAL KONCERTGÆSTER

Det samlede antal koncertgæster oplyst af arrangørerne, kan sammenholdes med 3's data, da det indeholder både danske og internationale gæster. De fleste kilder indikerer, at der deltog 40.000 i koncerten, selvom enkelte kun skriver 35.000. 3's data bør altså tilnærme sig dette antal i løbet af aftenen.

Figur 16: Antal koncertgæster fordelt på gæstetype



Figur 16 viser antallet af de forskellige gæstetyper i løbet af dagen. Endagsbesøgende er defineret som en person, der har opholdt sig i området mere end 90 minutter, men færre end 10 gange på en måned, og bør derfor udgøre klart de fleste koncertgæster. Antallet af endagsbesøgende i løbet af dagen underbygger, at de fleste koncertgæster udgøres af denne kategori – lav i løbet af dagen, stejlt stigende op til koncertens starttidspunkt med et kraftigt fald bagefter. For timerne omkring koncerten (19-23) ligger antallet for gæstetypen "Endagsbesøgende" dog noget lavere end det oplyste antal koncertgæster. Der er et maksimum på 28.000 kl. 20. No-shows trækker naturligt antallet ned, men kan næppe forklare hele forskellen. Estimatet på 28.000 udgør kun 80% af 35.000 og 70% af 40.000.

Medregnes personer fra de andre gæstetyper nærmer tallet sig dog de 35-40.000. Sammenlægges alle kategorier bortset fra "beboer", rammer estimatet præcis 40.000 kl. 20. Noget kunne således tyde på, at koncertgæster også eksisterer i andre af kategorierne.

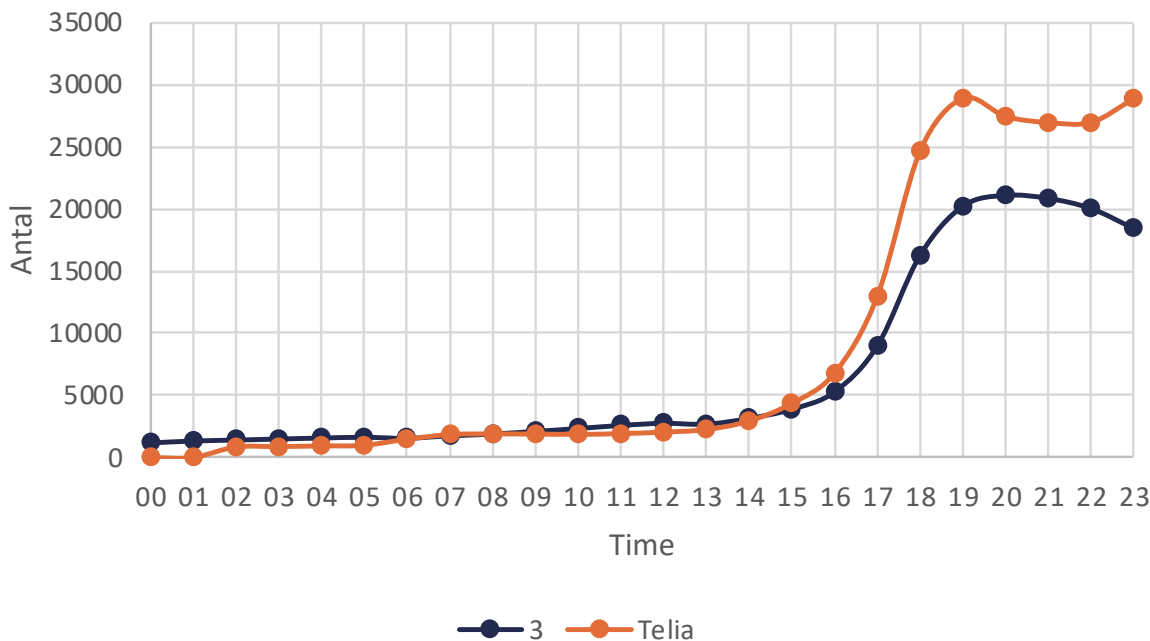
En kuriositet er dog det konstant høje niveau i "beboer"-kategorien, som viser omkring 7.500 personer først og sidst på dagen. 3 baserer antallet af beboere i et område, som rækker ud over POI'en, men som varierer baseret på mastepaceringer og celledækningen. Årsagen er, at "beboer"-kategorien kun kan fastsættes ved at kigge på et større udsnit af celler, som personen opholder sig i igennem længere tid. I denne proces vil celler, som personen kun opholder sig i kortvarigt, også inkluderes. Personer som bor tilstrækkeligt tæt på, vil derfor kunne blive inkluderet i "beboer"-kategorien, selvom de ikke reelt har opholdt sig i POI'en. 3 vurderer at området har en radius på 1000m omkring POI'en.

Fra figur 16 kan man konkludere, at udviklingen i gæsteantal virker validt for endagsbesøgende i løbet af dagen, men at det samlede antal er 70%-80% for lavt. Flere af de andre kategorier indeholder givetvis koncertgæster, men hvor mange er behæftet med stor usikkerhed. I den videre analyse benyttes derfor kun endagsbesøgende fra 3, da denne kategori kan antages alene at indeholde koncertgæster.

FORDELING AF GÆSTER

Sortes de udenlandske gæster fra i 3's data, kan 3's og Telias data sammenlignes for hele d. 22. juni. De udenlandske gæster udgjorde lidt over 6.700 af de registrerede endagsbesøgende i 3's data, og er fratrukket i nedenstående tabeller for at understøtte sammenligningen i antal danske gæster.

Figur 17: Antal danske koncertgæster

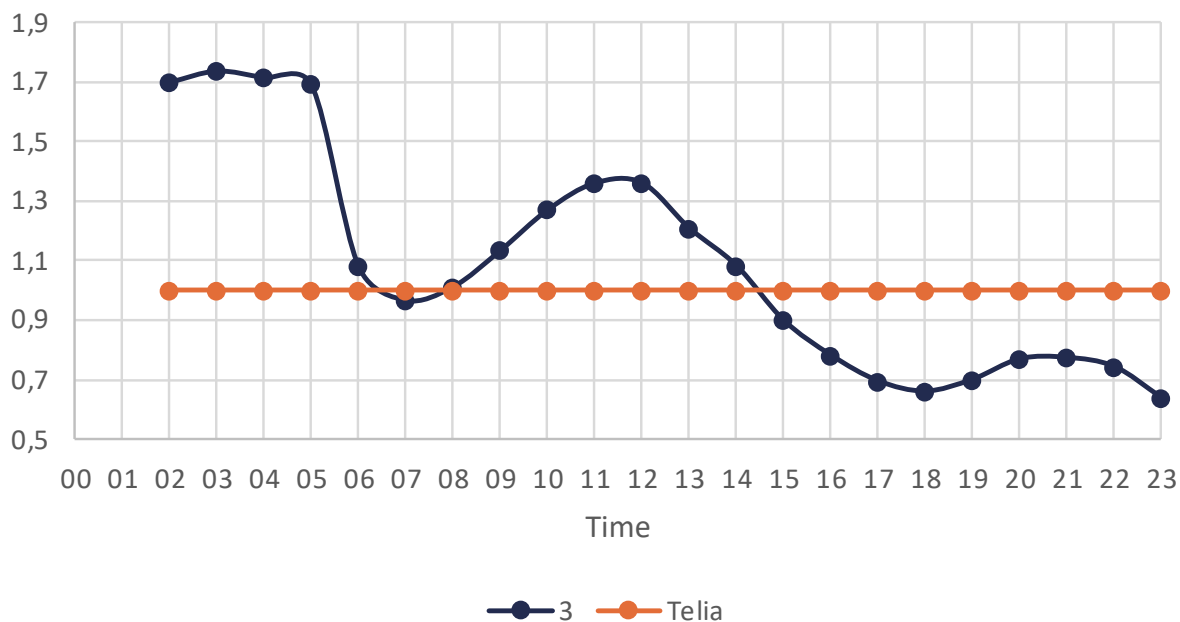


Figur 17 viser, at tallene følger hinanden relativt pænt i løbet af dagen, og at de stemmer godt overens med besøgs mønsteret til et større event. Telia viser små 30.000 gæster på koncerttidspunktet, hvilket passer godt, da man foruden danske gæster estimerede cirka 10.000 tyske gæster. Samlet nærmer antallet sig altså 40.000 for Telia, hvilket tyder på, at Telia faktisk estimerer det samlede antal mere præcist end 3. 3's antal af danske gæster er betydeligt lavere på kun lidt over 20.000 gæster kl. 20.

Derudover viser Telias data en bølgedal under koncerten (timerne 20-22), som også ville opstå i 3's data, såfremt gæstetypen "in-transit" var tilføjet ("in-transit" viser en markant bølgedal på koncerttidspunktet). Forklaringen herpå kunne være, at en del personer er beskæftiget i koncertområdet umiddelbart før og efter koncerten med sikkerhed og logistik.

Der eksisterer altså en uoverensstemmelse mellem 3 og Telia: En mulighed er, at der er andre gæstetyper end endagsbesøgende hos 3, som har deltaget i koncerten. Dette er sandsynligt, da 3 ellers markant underestimerer antallet af koncertgæster. Alternativt indeholder Telias data for mange personer, men eftersom Telias estimerer passer pænt med de officielle tal, tyder det ikke på at være tilfældet.

For at tydeliggøre forskellen i estimering af danske gæster, er Telia og 3's data indekseret i figur 18. Telias værdier indekseret til 1, og viser derfor hvordan 3's tal udvikler sig relativt til Telias tal i løbet af koncertdagen.

Figur 18: Danske koncertgæster indekseret (Telia = 1)

Umiddelbart er koncertgæsternes antal ikke estimeret særlig konsistent mellem teleselskaberne, da det skiftevis under- og overestimeres. F.eks. estimerer 3 i gennemsnit hele 26% færre gæster end Telia i tidsperioden 15-23, som er de timer, hvor koncertgæster reelt har opholdt sig i området. Selvom tendensen i gæsteanstal så pænt ud i figur 17, dækker det altså over temmelig store relative forskelle.

Der er også foretaget en korrelationsanalyse af datapunkterne, som viser en høj korrelation. Det betyder, at høje og lave værdier for det meste følger hinanden pænt mellem 3 og Telia. Datapunkterne er dog samlet i to klynger for henholdsvis lave og høje værdier, hvilket gør en korrelationsanalyse uhensigtsmæssig, og den er derfor ikke medtaget her.

En umiddelbar konklusion på KPI'en "antal gæster" er, at der eksisterer en betydelig forskel i 3's og Telias data, men at forskellen givetvis kan tilskrives opdelingen af gæstetyper i 3's data, hvor koncertgæsterne er opdelt i flere forskellige kategorier. Til gengæld er udviklingen i gæsteanstal i løbet af dagen nogenlunde konsistent mellem datakilderne, selvom 3 skiftevis ligger over og under Telias antal. Grundet usikkerheden i estimeringen af "beboer"-kategorien hos 3, har Telia i denne case umiddelbart det bedst estimat for antal gæster.

HVOR KOMMER GÆSTERNE FRA

Både 3's og Telias data angiver, hvilke kommuner de danske gæster kommer fra, og i nedenstående analyse vises de største kommuner samt aggregeret data på regionsniveau. 3's data vises kun for kategorien "endagsbesøgende".

Regioner

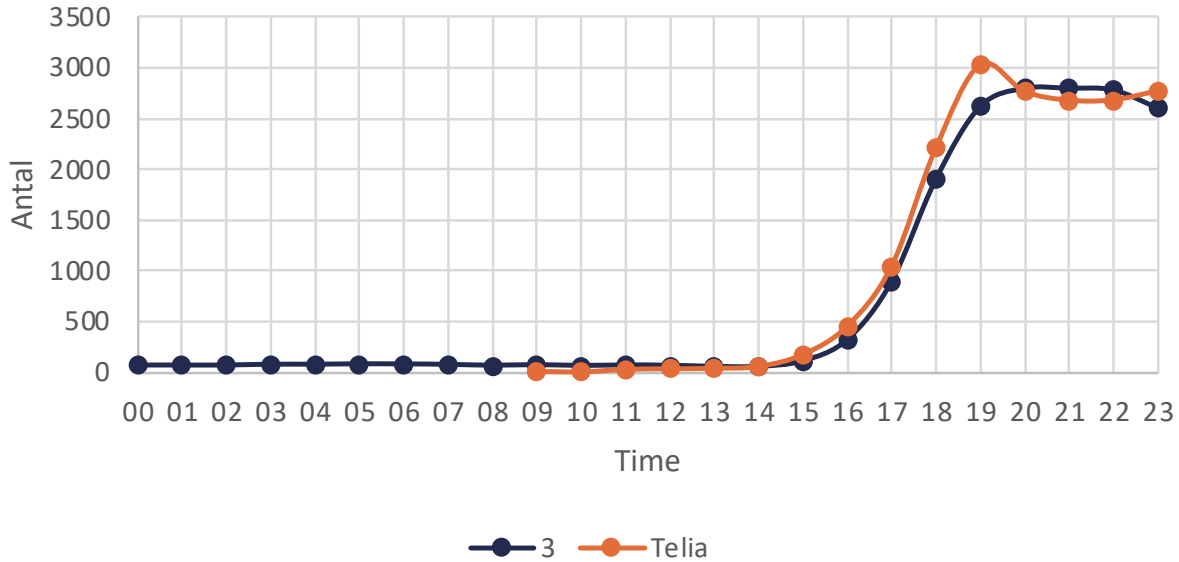
Opdeles koncertgæsterne på regionsniveau, gentages udviklingen generelt fra figur 17: Lave niveauer først på dagen og en kraftig stigning i timerne umiddelbart inden koncerten. Telia estimerer igen et højere antal gæster på selve koncerttidspunktet med peaks før og efter, men bemærk hvordan niveauerne er næsten ens for Region Hovedstaden og delvist Region Nordjylland. Derudover indeholder Telia en relativt stor andel koncertgæster uden lokationsangivelse (NA).



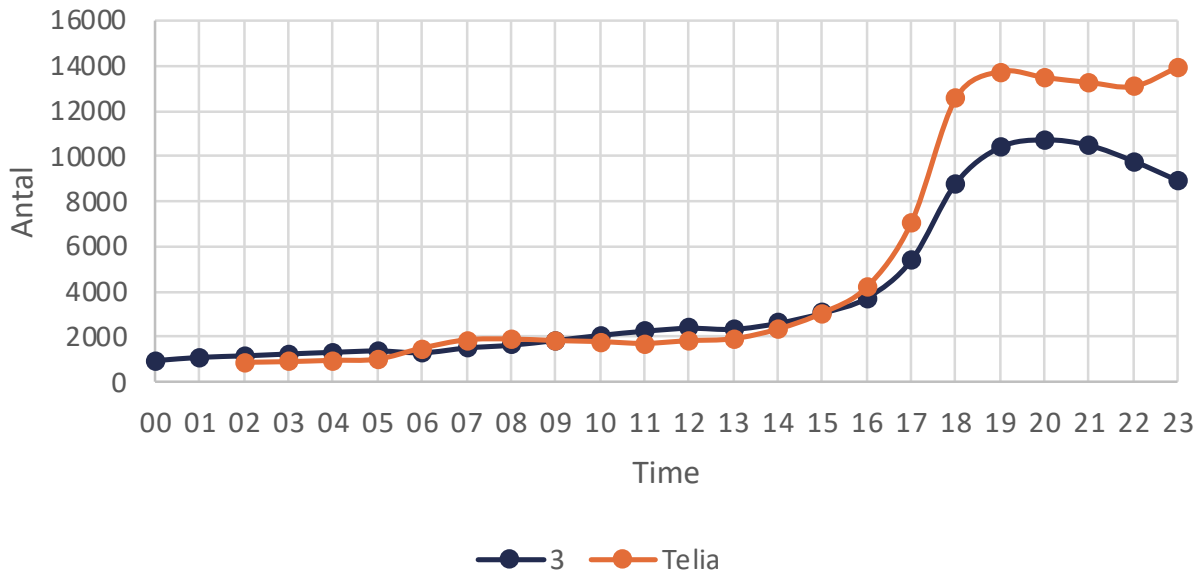
Foto: Daniel Villadsen

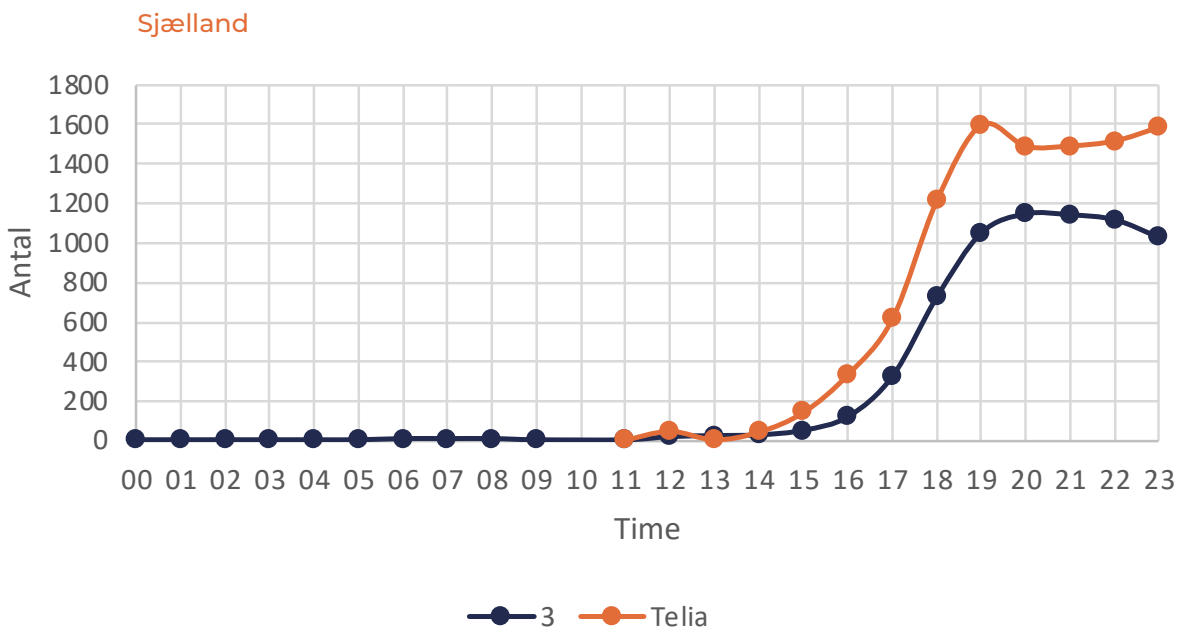
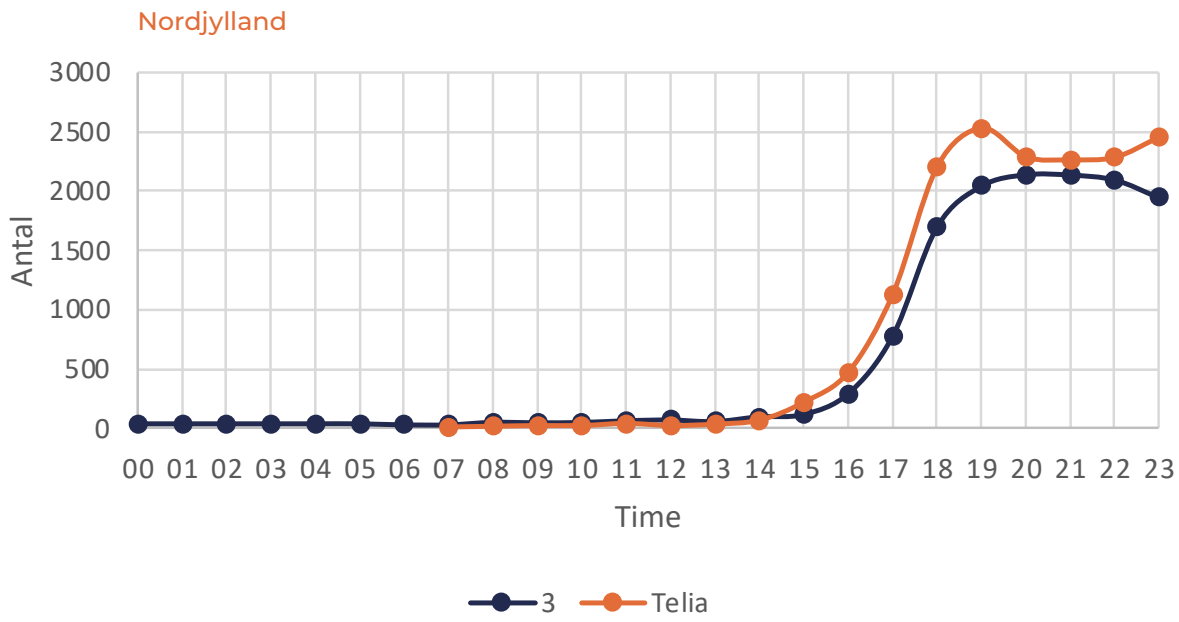
Figur 19-24: Antal koncertgæster fordelt på regioner

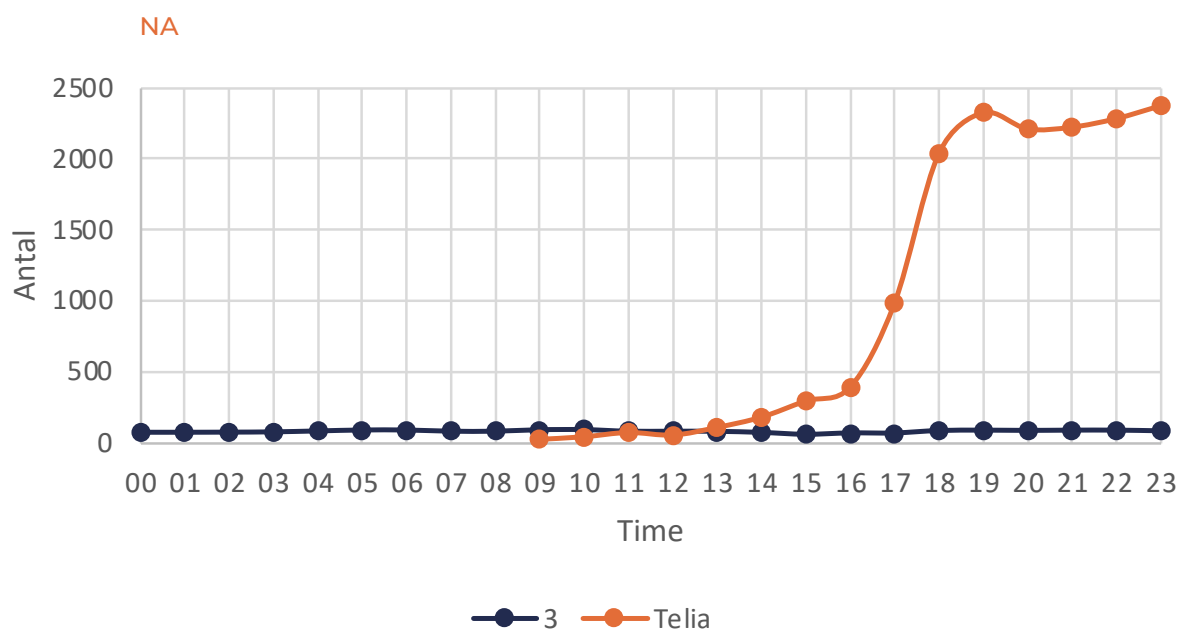
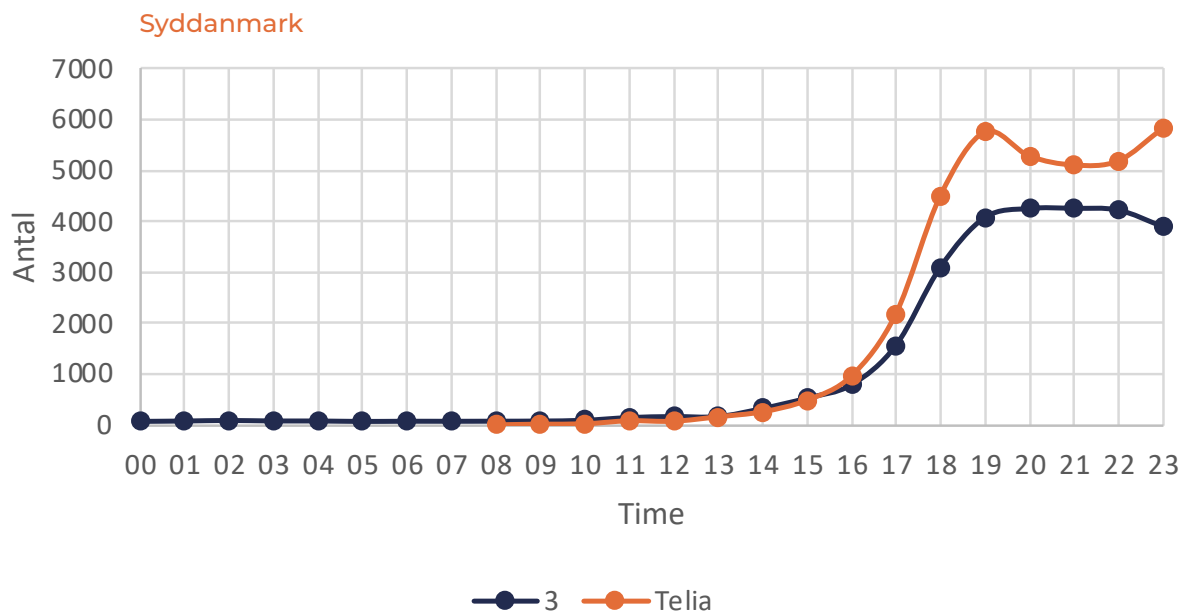
Hovedstaden



Midtjylland







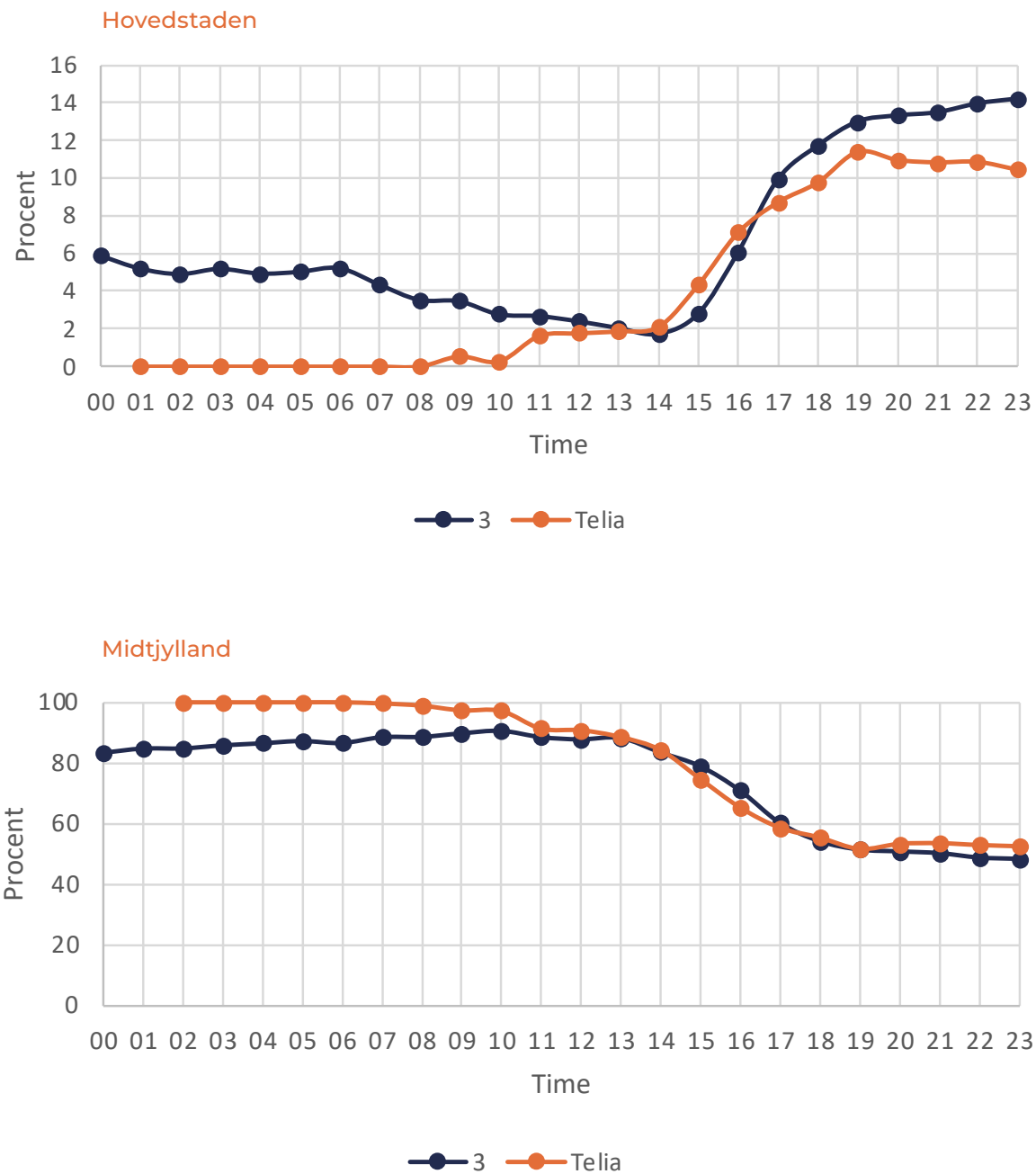
Det er svært at konkludere på selve antallet af koncertgæster fra de forskellige regioner, når 3 og Telia ikke viser ens tal. En bedre indikator er den relative fordeling af gæster, som bør være ens mellem 3 og Telia, hvilket illustreres i figurerne nedenfor – andelen af gæster fra de forskellige regioner i løbet af dagen.

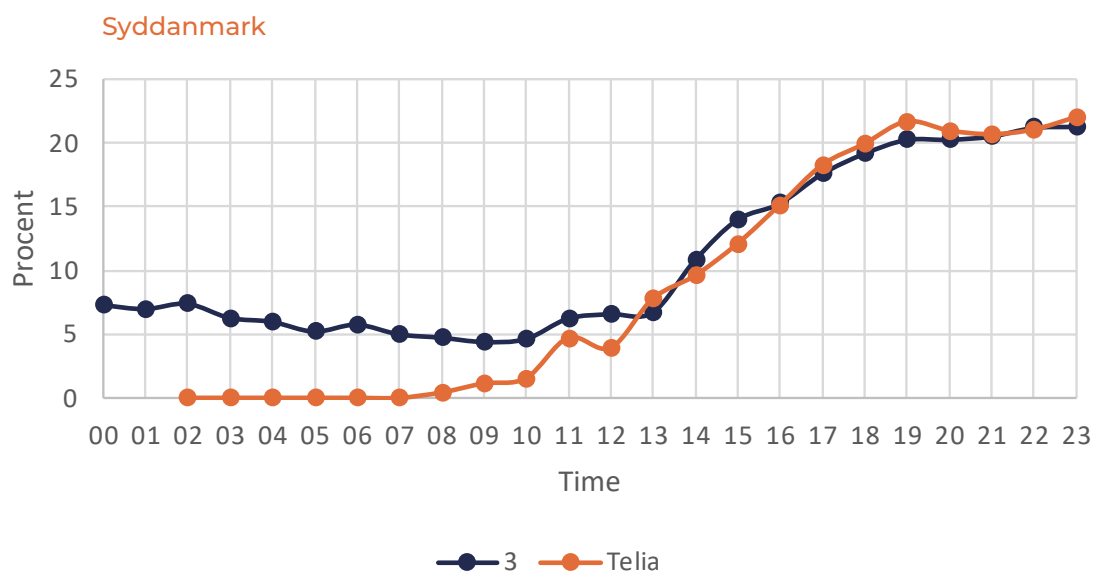
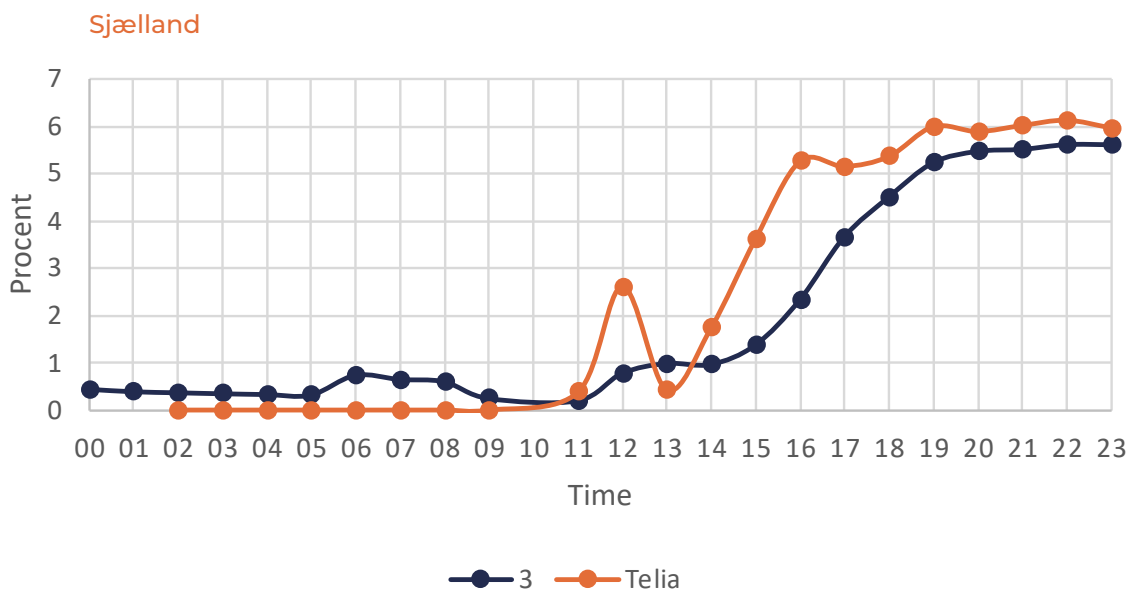
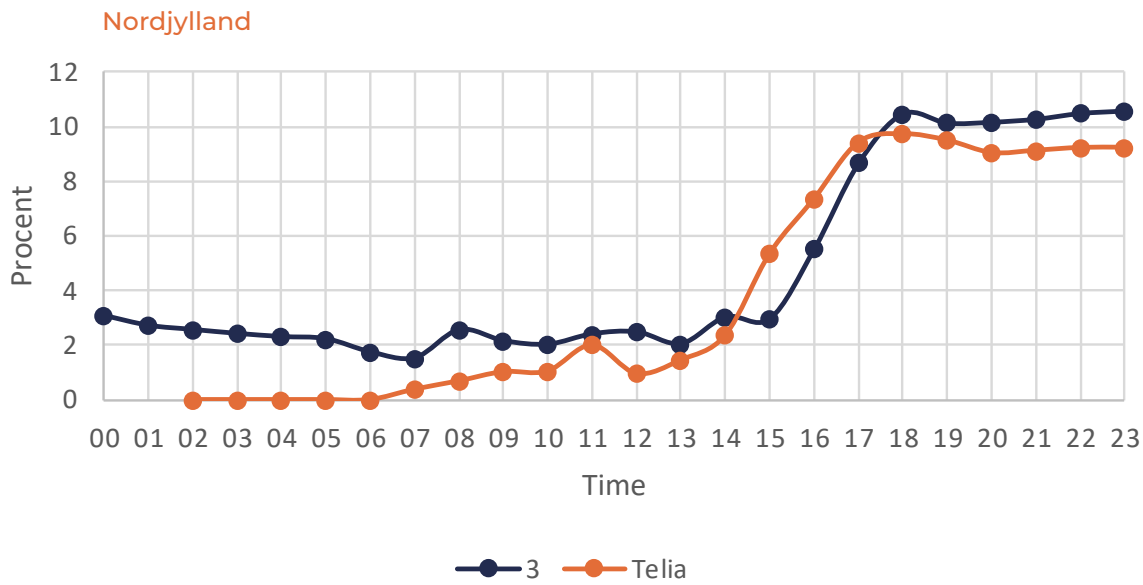
For første halvdel af dagen udgør personer fra Region Midtjylland næsten alle i området, imens tallet falder kraftigt i timerne op til koncerten og stabiliseres på omkring 50%. Dette virker sandsynligt, da personerne som både arbejder fast i området og arbejder isoleret i forbindelse med koncerten, kommer fra nærområdet. 3 viser dog en noget mindre andel end Telia her.

Fra kl. 15 og frem har 3 og Telia meget ens relative fordelinger på tværs af alle regioner. Forskellen for Region Sjælland, Region Syddanmark og Region Nordjylland er procentuelt meget små, imens de er lidt større for Region Hovedstaden og Region Midtjylland. Procentfordelingerne divergerer altså i et mindre omfang, men overordnet viser data et rimelig konsistent billede.

Der kan altså gives et ganske kvalificeret svar på, hvor i landet koncertgæsterne kommer fra. For det eksakte antal gæster fra hver enkel region, er de samme problematikker gældende, som diskuteret tidligere.

Figur 25-29: Procentandel gæster fordelt på regioner





Kommune

Analyserne er også foretaget på kommunalt niveau, som af pladshensyn ikke er vist her. Det generelle billede er, at kommuner med mange koncertgæster generelt viser konsistente tal mellem 3 og Telia, men for kommuner med få koncertgæster kan der opstå temmelig store usikkerheder. En generel konklusion er derfor, at usikkerheden i teledata er størst for mindre datagrupperinger. En tommelfingerregel i denne case kunne tyde på, at grupperinger på mindre end 100-150 personer er behæftet med særlig stor usikkerhed.

3.4.3 CASE-KONKLUSION

KPI'en "antal personer" kan ikke entydigt besvares, da Telia generelt viser et højere antal end 3. Gæstetype-variablen hos 3 er givetvis årsagen. Formodningen er dog, at Telias data rammer gæsteantallet ganske godt, og at 3's data kan gøre det samme, hvis variablen for gæstetype bliver udlagt mere præcist. Det relative antal gæster, som indikerer, hvornår der er henholdsvis mange og få gæster i området, kan derimod besvares rimelig præcist. Det samme gør sig gældende for KPI'en, "hvor gæsterne kommer fra": Det eksakte antal gæster er usikkert, men den relative fordeling er nogenlunde valid – det gør sig gældende for både regions- og kommuneniveau – på nær kommuner med færre end 100-150 koncertgæster. Alt i alt giver både Telia og 3 nogle gode indikationer af gæstetrykket, dog er begge teledatakilder lavere end forventet i absolutte tal. Telias lavere andel kan formentligt tilskrives de "manglende" internationale gæster i data.



Foto: Nordsjælland, Niclas Jessen

KONKLUSION



4

På tværs af rapportens cases (beskrevet på side 5), kan det konkluderes, at teledata i en vis udstrækning kan sige noget om den samlede gæstemængde. De absolutte tal kan i nogle tilfælde, hvor andelen af internationale gæster er lav, mastedækningen god og den samlede gæstemængde høj nok, anvendes. Der er dog, som forbeholdene også indikerer, en række forhold, der skal være stabile, før teledata kan give et tilfredsstillende billede i absolutte tal.

Overordnet set kan det derfor også konkluderes, at det – i flertallet af de cases projektet har beskæftiget sig med – ikke er muligt at give et kvalificeret bud på antallet af gæster, og derfor ej heller antallet af endagsturister i denne rapport.

Mængden af endagsturister kan samlet set ikke valideres tilstrækkeligt og rammen for, hvordan endagsturismen bør håndteres eller tolkes som absolut tal i teledata, kan derfor ikke defineres. Der hersker dog ingen tvivl om, at teledata har et udviklingspotentiale, der på sigt kunne give et kvalificeret bud på antallet af danske endagsturister i områder, hvor datakvaliteten tillader det (antal observationer og mastedækning).



Foto: Læso, Ditte Ingemann

Potentialet i teledata ligger i, at begge teledataselskaber i flere cases udviser gode trends i data: Data viser således de rigtige tendenser, om end det absolutte tal ikke stemmer, hvilket allerede nu giver teledata sin berettigelse fx til estimering af gæstetryk i et område, hvor mastedækningen er god og antallet af observationer højt nok. Samtidigt muliggør retvisende tendenser i data, at et mere præcist absolut tal for (endags)turisme kan defineres.

Endeligt er der nogle væsentlige konklusioner for hver af teleselskaberne:

Telia performer – på trods af de ”manglende” internationale gæster – flot på blandt andet Rammstein-, Egeskov- og København-casen. De rammer trends og delvist absolutte tal, hvormed de repræsenterer en relativt sikker datakilde, med forbeholdet for, at der i flere tilfælde vil ske en underestimering af det samlede antal gæster, da de internationale gæster ikke ekstrapoleres. Selvsamme er Telia meget eksplicit omkring.

Inddelingen i 3's gæstetyper har en række fordele, da det indikerer andelen overnattende gæster, beboere, mm., som Telia ikke har mulighed for i deres data. Med kategoriseringen følger dog nogle forbehold, da flere kategorier er behæftet med usikkerhed – særligt når der arbejdes med mindre afgrænsede områder.

I tillæg kræver kategoriseringen af beboere et særligt vågent øje, da beboerne er defineret i større områder, end den POI (Point of Interest) der udvælges. Sagt med andre ord: Arbejdes der med et særligt afgrænset område, som i denne rapport, vil beboer-kategorien altid fange beboere i X-radius udenfor undersøgelsesområdet. Beboertallene kan derfor ikke antages som værende valide, hvis man forventer, de fortæller noget om, hvor mange beboere der deltager / gæster undersøgelsesområdet.

Samlet set performer 3 bedst i Egeskov- og Rammstein-casen. De rammer trends og i enkelte tilfælde absolutte tal, ofte ses der dog nogle bemærkelsesværdige overestimeringer, og de repræsenterer således en datakilde, der skal bruges med forbehold for dette.



HVORDAN SKAL MAN ARBEJDE MED TELEDATA ("HOW-TO-GUIDE")

5



I denne "How-To-Guide" opridses en række overvejelser eller trin, der med stor sandsynlighed kan forbedre kvaliteten af de data, du tilegner dig eller arbejder med. For når du skal arbejde med teledata, er det vigtigt at gøre sig en række overvejelser, der sikrer, man styrer rundt om kendte faldgruber såsom passende mastedækning og observationer, mangel på alternative valideringskilder og metodiske finurligheder i teledatastrukturen.

Mange af punkterne vil du også blive vejledt i af teledataudbyderen, hvis du påtænker at købe data til en specifik case – men du kan med fordel selv gøre forarbejdet og vurdere, om teledata overhovedet er det rette valg til din case, og hvis det er, sikre bedst mulig kvalitet og validitet i dit dataudtræk.

Forhåbningen med denne guide er desuden at skabe fælles forståelse og anvendelse af teledata, så vi sammen kan øge kvaliteten af data – og vores viden – om endagsturismen, sikre ensartede målemetoder og ikke mindst fortolkninger af teledata i dansk turisme.

5.1 TELEDATAUDBYDERE

Der er en lang række forskelle mellem de to teledataudbydere Telia og 3, der kan have en påvirkning på, hvilken udbyder der bedst egner sig til en specifik case. For så vidt muligt at hjælpe valget på vej, er der i dette afsnit lagt vægt på nogle af de mest centrale forskelle. Først gennemgås grundlæggende forskelle i, hvordan 3 og Telias data er opbygget, dernæst kigger vi på de gæstetyper, der er defineret i 3's data og endeligt på andre vigtige forskelle og overvejelser, du bør gøre dig, hvis du vil arbejde med teledata.



5.1.1 GRUNDLÆGGENDE FORSKELLE I DATASTRUKTUREN

	TELIA	3
ID	Telia opererer pt med et 24-timers ID. Det betyder, at det givne ID randomiseres hver nat, og kan derfor ikke følges over flere dage*	3 opererer med et måneds-ID. Det betyder, at det givne ID randomiseres i slutningen af hver måned, og kan derfor følges over flere dage* Det betyder også, at 3 opererer med en række gæstedefinitioner. Disse er uddybet i Afsnit 5.1.2.
Nationalitet	Telia har både danske og internationale roaminger fra hele verden (internationale ekstrapoleres** ikke)	3 har både danske telefoner og internationale roaminger fra hele verden (internationale ekstrapoleres**)
Repræsentativitet	Telia ekstrapolerer data ud fra deres markedsandel, og repræsenterer derfor en beregnet andel af hele den danske befolkning og deres aktiviteter. Telia ekstrapolerer ikke internationale roamere, men stiller gerne den ikke-ekstrapolerede data til rådighed.	3 ekstrapolerer data ud fra deres danske markedsandel, og repræsenterer derfor en beregnet andel af hele befolkningen og deres aktiviteter. 3 arbejder også med demografiske data i de deres ekstrapolering af danskere. 3 ekstrapolerer data på internationale roamere, ud fra andelen og antallet af roamere fra et givent internationalt netværk, der anvender 3's netværk. Målingen af roamere på deres netværk bruges bl.a. til at estimere og derfra opjustere underrepræsenterede roamere.

*Alle data er GDPR-compliant. Teledata indeholder ingen personhenførbare data, men viser bevægelsestrends. Der skal foretages minimum 5 unikke rejser, før de vises i data.

**Ekstrapolering er en beregning, der skaber ukendte værdier ud fra kendte værdier eller sammenhænge. Fx bruges demografiske informationer om den danske population ofte til at op- eller nedregne surveydata i en undersøgelse, så man antageligt kan sige noget om et repræsentativt udsnit af den danske befolkning. Metoden er behæftet med en række usikkerheder. For mere information om hhv. Telias og 3's opregninger, bedes du kontakte teleselskaberne direkte.



Foto: Freja, Niclas Jessen

5.1.2 3'S GÆSTEDEFINITIONER

Tilbage i 2021 blev det første større udviklingsprojekt med teledata gennemført med hjælp fra genstartsmidler fra Erhvervsministeriet. Her blev teledata defineret og kategoriseret ved teleselskabet 3 ud fra antallet af og længden på besøg i et givent område. Definitionen kan kun laves og bruges i 3's data, fordi de har et måneds-ID, hvorimod Telia pt opererer med et døgn-ID:

BESØGSTYPE	
BEBOER	Defineres ud fra ens "hjemme-zone", der beregnes på baggrund af "inferred home cells analytics" (baseres på bl.a. inaktivitet om natten)
PENDLER	Er ikke klassificeret som beboer i området, bliver i området i mindst 90 min., og har flere end 10 besøg i området i måneden.
I TRANSIT	Bevæger sig gennem et område, opholdet er under 90 min.
ENDAGS TURISTER	Bliver i området i mindst 90 min. Har færre end 10 besøg i måneden.
KORT OPHOLD	Overnatter 1-3 nætter, og overnatter mindre end 22 nætter i det givne område (kan opdeles i 1, 2 og 3 nætter).
LANGT OPHOLD	Overnatter i 4+ nætter, og mindre end 22 nætter i det givne område (kan opdeles i 4-6 nætter, 7-13 nætter og 13+ nætter).

5.2 UDVÆLGELSE AF CASE/ VURDERING AF DATA

Inden du giver dig i kast med at bruge teledata til at sige noget om antal gæster eller typen af gæster i et givent område er, skal du overveje, om din case egner sig til at bruge teledata. I nogle tilfælde er en case givet – og så kan du bruge guiden til at vurdere, om teledata egner sig. Det kan også være, du allerede har købt teledata, eller påtænker at bruge det teledata, som bliver tilgængeligt i den nye nationale dataplatform for turisme. Så kan guiden hjælpe dig med at afklare, hvor gode de data, du har tilegnet dig, faktisk er.

5.2.1 OMRÅDE, MASTERDÆKNING & FORVENTET ANTAL GÆSTER

Først om fremmest skal du, når du vælger eller vurderer en case, overveje om din case ligger i et område, der egner sig til at bruge teledata. I rapporten her, blev tre forskellige niveauer defineret, da hvert niveau (afhængigt af casen) giver grobund for forskellige spørgsmål og overvejelser. De tre niveauer kan groft opdeles i to grupper, hvor du kan stille dig selv følgende spørgsmål:

	BESKRIVELSE	SPØRGSMÅL
NIVEAU 1	Lille afgrænset geografisk område (fx koncert, små events, attraktion, stadion)	<ul style="list-style-type: none"> • Er case-området tydeligt afgrænset fra andre områder med nogen eller høj aktivitet (virksomheder, skoler, detail og indkøbsmuligheder, cafeer, restauranter, mm.)?
NIVEAU 2	Større afgrænset geografisk område (fx feriecenter/-park, feriehusområde, store events)	<ul style="list-style-type: none"> • Hvordan placerer området sig i forhold til beboelse? • Hvor mange gæster forventes at besøge området?
NIVEAU 3	Byer eller en kommune	<ul style="list-style-type: none"> • Er der i by- eller kommunegrænserne områder med nogen eller høj aktivitet, der kan påvirke – eller er vigtige for – mine data? • Er der i kommunen en eller flere større rastepladser (fx kommuner med motorvej), som kan påvirke mine data? • Inkluderer tidsperioden for mit dataudtræk ferieperioder, der kan påvirke fx pendler-/endagsturismendelen (+/- 10 besøg)? • Har jeg nogle gode valideringskilder til rådighed, der kan hjælpe med at vurdere datakvaliteten på så stort et område?

Derudover kan du med fordel tjekke www.mastedatabasen.dk for at se, hvor mange master der ligger i netop det område, du gerne vil undersøge. Antallet af master i nærområdet har nemlig en betydelig påvirkning på, hvor akkurate dine data er.

Eksempel Da dette projekt udvalgte cases, blev flere større feriehusområder udelukket på grund af enten dårlig mastedækning eller afgrænsning. I nogle tilfælde på grund af "støj" fra nærliggende områder med helårsboliger, indkøbsmuligheder eller mindre byer, som ikke lå langt nok fra det case-område, vi overvejede at bruge. Helt konkret, kunne vi ikke isolere fx Landal Søhøjlandet fra den nærliggende by Gjern. Dels på grund af antallet af master, dels på grund af antallet af gæster i området. Landal Seawest lykkedes alene, fordi vi valgte at inkludere hele det omkringliggende feriehusområde.



5.2.2 BASELINE

I arbejdet med teledata bør du altid sammenligne dine data med en baseline. Det betyder, at når du arbejder med et givent område i en given tidsperiode, fx en festival, også bør tilkøbe data fra det samme område i en periode, hvor der ikke foregår nogen særlige aktiviteter. På den måde får du en god idé om, hvor meget aktivitet der er i området normalt, som du kan sammenligne direkte med den periode, hvor du har en særlig case, du vil undersøge.

Det ville i dette projekt fx have været en fordel i Rammstein-casen, hvor der var et markant højere antal observationer "i transit" og "beboere", end vi havde antaget givet det område, vi havde udtrukket. Det udelukkede muligheden for at analysere, om – og i så fald hvor stor en andel – der normalvis registreres som beboere eller i transit, og om andelen steg eller ændrede sig på grund af koncerten i området.

5.2.3 REALISTISKE FORVENTNINGER

Sæt realistiske forventninger til hvilke svar og tal, du forventer teledata kan levere. Start altid med at validere det samlede antal mennesker i området – er tallet eller tendenserne realistiske? Stemmer det overens med andre data fra området, hvis du har nogen? Først når du har valideret det samlede tal, kan du gå i dybden med andre tal eller beregninger. Den naturlige proces, hvis du skal arbejde med endagsturister kunne se ud på følgende måde:

1. **Undersøg antal gæster**
 - a. Validér mod andre data
 - b. Sammenlign med en baseline

Når/hvis du stoler på det samlede antal, kan du arbejde i dybden med KPI'er som:

2. Hvor kommer gæsterne fra?
3. Hvor længe opholder de sig i området?
4. Hvor stort er gæstetrykket (densitet/intensitet)?
5. Eventuelle maxbelastninger af et område – eller kortlægning af spidsbelastningsperioder (dage/tidsperioder)

Arbejder du med endagsturisme ud fra et område (fx gæster i København, der rejser til Roskilde, Helsingør, mm. – eller hvor feriehushusgæster bevæger sig hen, når de forlader feriehusområdet), kan du med fordel kigge på:

6. Hvilke kommuner/byer rejser gæsterne til (grupperet over en periode)
7. Hvilke tidspunkter forlader gæsterne typisk området – og kommer tilbage igen?
8. Hvor stor en del af gæsterne bliver i området?

Alle spørgsmål/KPI'er fra 2-8 afhænger dog af, at du både har valideret det samlede antal gæster, og at du har en stor nok gæstemængde til at kunne se noget retvisende i data (uden at skulle gruppere data i for store og uhensigtsmæssige grupper).

5.3 STYRKER OG SVAGHEDER VED TELEDATA

I forlængelse af afsnittene ovenfor, vil dette afsnit påpege relevante forskelle og fund fra selve analyserne af rapportens cases, du med fordel kan notere dig, hvis du vil arbejde med teledata:

5.3.1 EKSTRAPOLERING

Vores samlede konklusion peger på, at der i flere tilfælde sker en overestimering af danske gæster hos 3 og en underestimering hos Telia, når der er tale om mindre områder med nogenlunde mastedækning og gæsteflow. Dette kan være forårsaget af forskellige metodiske tilgange til, hvordan hhv. 3 og Telia opregner (ekstrapolerer*) deres data på danske gæster, som naturligt har en direkte påvirkning på det samlede antal gæster teleselskaberne hver især estimerer.

Ekstrapoleringen anses om en forretningshemmelighed, og vi har derfor ikke mulighed for konkret at vurdere, hvor stor usikkerhed data er forbundet med i forskellige scenarier eller cases. Vi kan blot påpege, at mængden af mennesker, den fysiske afgrænsning og mastedækningen er helt centrale for validiteten.

5.3.2 "BEBOER"-KATEGORIEN (3)

"Beboer"-kategorien i 3's data, beregnes på baggrund af "inferred home cell analytics", hvor trends over tid (bl.a. inaktivitet om natten) definerer det enkelte signals "hjemmezone". Hjemmezonen kan dog overlappe med den POI (case-området), der arbejdes med, og det er derfor usikkert, hvorvidt en beboer, som har en nærliggende hjemmezone, faktisk opholder sig i POI'et – eller bare er hjemme. Uanset længden på opholdet, vil signalet, hvis det overlapper med hjemmezonen, altid registreres i kategorien "beboer".

Dette er i sig selv en ganske udmærket logik. Der opstår bare en problematik med at angive endagsturister – eller gæster til et event – hvis beboere fra nærområdet automatisk "ekskluderes" som endagsturister.

Eksempel I Rammstein-casen blev der i en radius af 1 km omkring Ceres Park Arena medtaget beboere i data. Det var således registreres over 7.000 beboere "i" Ceres Park Arena på koncerttidspunktet. Det er dog usikkert hvorvidt – eller hvor mange – der deltog i koncerten, fordi deres hjemmezonen overlapper med det område vi har tegnet for Ceres Park Arena (se afsnit 3.4). Det er naturligvis problematisk, hvis en case ligger i et område med beboere, der også forventes at deltage i det event (el. lign), der undersøges. For kategoriseringen i teledata kan ikke vurdere, hvor stor en andel af de lokale, der faktisk deltager til eventet. På den anden side er man sikker på, at alle øvrige registreringer i kategorien endagsbesøgende, faktisk *er* endagsturister og har deltaget. Det vil i sagens natur (igen) afhænge af casens fysiske afgrænsning, hvor sikkert tallet bliver – Ceres Park Arena var ikke, trods sin lokation, afgrænset godt nok.

5.3.3 OVERNATTENDE GÆSTER

En relativ forskel i teleselskabernes datasæt er mængden af tid, som de kan følge de enkelte brugere i. Hvert teleselskab tildeler de enkelte brugere på deres netværk et unikt ID, som teleselskabet herefter kan følge hen over en fastsat tidsperiode med henblik på at analysere f.eks. bevægelsesmønstre. Der er dog begrænsninger på hvor lang tid teleselskaberne må følge de enkelte brugere, før denne data skal slettes jf. GDPR-reguleringer. På nuværende tidspunkt kan 3 følge de enkelte brugere i op til en måned, mens Telia udelukkende følger brugerne i en enkelt dag. Dette, kombineret med det faktum at 3 opdeler deres brugere i mange flere kategorier (beboere, kortidsbesøgende etc.) end Telia, skaber meget højere kompleksitet i 3s datasæt, hvilket kan mudre det billede som datasættet skaber af bevægelsesmønstre i et givent område og gør det sværere at validere data, sammenlignet med Telia, som udelukkende følger de enkelte brugere i en enkelt dag og ikke opdeler deres brugere i underkategorier baseret på besøgstype (beboere, kortidsbesøgende etc.). Telias data virker derfor mere overskuelige og er mere "ligetil" at tilgå og arbejde med, mens 3s data dog potentielt indeholder mere information om bevægelsesmønstre og opholdslængder på tværs af dage og dermed potentielt kan give mere værdi afhængig af problemstillingen i et givent projekt. På samme tid vil det dog også kræve flere ressourcer (tid og kompetencer) at gennemarbejde og validere 3s data end det vil for Telias data.

5.3.4 ANTAL OBSERVATIONER

Antallet af observationer er afgørende – antallet af gæster i området skal være højt nok, før det giver mening at bruge teledata. Det kan være utrolig svært at vurdere, hvornår der er observationer nok, men ud fra den erfaring vi har draget i dette projekt, kan selv cases med forventeligt 1.000-1.500 mennesker, der ligger i gode områder med nogenlunde mastedækning, vise sig at være svære for teledataselskaberne at estimere med en tilfredsstillende sikkerhed.

Eksempel To gode modsætninger i denne rapport, hvis man vil kigge på antallet af mennesker pr. dag, er mellem Landal Søhøjlandet og Rammstein koncerten. Til koncerten kunne vi arbejde med antallet af gæster pr. time, fordi andelen af observationer var høj og mastedækningen god. I Landal Seawest er antallet af gæster på dagsniveau i teledata, efter vores vurdering, ikke retvisende. Antageligt fordi antallet af observationer er for lavt, og derfor over- eller underestimeres i ekstrapoleringerne i data.

Sagt med andre ord; Cases med mange observationer fungerer rigtig godt – men cases med få observationer har mange faldgruber, der skal tages højde for, før data kan bruges til at sige noget om det faktisk antal endagsturister eller gæster.

5.3.5 FAKTISKE TAL – ELLER TRENDS?

I forlængelse af "Antal observationer" er det relevant at understrege, at teledata i rigtig mange tilfælde godt kan sige noget om trends. Teledata viser de forventede udsving, der bør være (højere andel pendlere fra morgenstunden og om eftermiddagen, flere gæster i Landal og Egeskov i weekenderne udenfor højsæsonen og ved events som i Rammstein eller Heartland festival i Egeskov). Antallet af gæster er bare ikke retvisende i alle tilfælde, det estimerede antal gæster bør derfor ikke anvendes eller rapporteres ukritisk.

5.4 OPSAMLING

Teledata har en lang række potentialer og muligheder, men det kræver arbejde af fortolkeren, at få retvisende og brugbare tal og tendenser. Teledata er ikke noget "quick-fix", men har store potentialer – og fortsatte udviklingsmuligheder. Det kan give gode indikationer og (i nogle tilfælde) faktiske tal, hvis en god case er udvalgt.

På sigt kan mobilitetsdata hjælpe destinationer med at få den nødvendige viden om volumen af gæster og lokale, samt deres mobilitet, på de enkelte destinationer. Retvisende data om (endags)turisters bevægelse vil kunne hjælpe destinationer med at udvikle lokale spredningsstrategier og sikre en mere bæredygtig turisme.

