

Udvikling af statistikcenter	Ansvarlig	AXH
	Oprettet	25-11-2017
Projekt: 7464, Digitale relationer og datadreven informationsformidling	Side	1 af 12

## Statistikcenter

Notatet er udarbejdet i arbejdsplan 1 omhandlende "Data og viden" i projektet: *Digitale relationer og datadreven informationsformidling*.

Der er i projektet foretaget en række implementeringer af statistiske visninger i MicroSoft PowerBI. Visningerne har været rettet mod formål, der skal sætte fokus på videnformidlingen til landmænd. Notatet beskriver hvordan implementeringen er forløbet.

## Baggrund

Brugen af hjemmesider i videnformidling har været i hastig udvikling siden årtusindskiftet. Brugere ser i stadig stigende grad det som en selvfølge, at de kan finde information hurtigt og nemt – uden at de nødvendigvis behøver at komme i direkte kontakt med afsenderen. Tidligere havde man som afsender et klart indblik i f.eks. hvor mange opringninger eller mails man modtog, og hvor store oplag af trykte materialer, der kunne afsættes.

Med introduktionen af hjemmesider, mobilapplikationer m.v. benytter brugerne i stadig stigende grad sig af selvbetjening. Dette nødvendiggør, at man på andre måder end tidligere, måler på brugernes anvendelse af viden – og tilfredshed med informationsformidlingen generelt. For at gøre det muligt at optimere formidlingen, er man selvsagt nødt til også at have indblik i, hvilken information det er brugerne efterspørger.

Almindelig webstatistik har været anvendt i Landbrug & Fødevarer i en lang årrække. Bl.a. har man set på data fra Google Analytics og udviklet egne applikationer til analyser af logdata. Men der viser sig et stadigt stigende behov for løsninger, der dels giver bedre muligheder for at samle statistik fra mange forskellige kilder, dels gør det muligt at få statistikkerne ud til en større gruppe af medarbejdere. Dermed øges sandsynligheden for at de tal, der er til rådighed, også reelt får indflydelse på den videnformidling, der foregår.

## Implementering

### Opstart

I 2017 har der været fokus på at finde frem til hvilken platform SEGES fremadrettet skal bruge til opbevaring og præsentation af statistiske data. Der er de seneste år sket en hastig udvikling indenfor området omkring datamanagement og datastorage. Bl.a. er kombinationen af Elasticsearch og Kibana blevet brugt bredt i huset.

Valget af platform blev foretaget på baggrund af følgende parametre:

#### 1) Integrationsmuligheder.

Det skulle være en løsning, der var i stand til integrere til mange forskellige kilder. Herunder bl.a.

- **Google Analytics:**  
Her opsamles alle oplysninger om trafik fra SEGES' hjemmesider.
- **SEGES' centrale brugerdatabase**

Brugerdatasen indeholder oplysninger om alle, der har en DLI-Ident og dermed kan logge på SEGES' løsninger. Herunder registreres brugernes type, f.eks. om de er landmænd, konsulenter, firmaabonnenter, studerende forskere osv. Landmænd kan selv opdatere deres oplysninger i brugerdatabasen og dermed f.eks. give information om deres produktion. Disse informationer kan så efterfølgende bruges til at målrette informationsformidlingen, så den bedst muligt matcher den individuelle bruger.

- **Conference Manager**

Dette system bruges til at administrere bl.a. SEGES' konferencer og møder. Nogle af disse kræver betaling, andre er gratis. Da kurser og møder er en vigtig del af SEGES' formidling, er det selvfølgelig vigtigt at have overblik over, om information når helt ud til landmændene af denne vej. Eller om der er brug for yderligere tiltag, f.eks. i form af artikler på hjemmesiderne.

- **Facebook**

Det bliver mere og mere almindeligt, at landmænd bruger Facebook i faglig sammenhæng. Bl.a. findes der en dansk gruppe for Malkekvæg med mere end 4.000 medlemmer. Mange landmænd opretter også åbne eller lukkede grupper på Facebook for deres gård og bruger den som Intranet for bedriften. Omtalt bl.a. i LandbrugsAvisen: <https://landbrugsavisen.dk/ekspert-facebook-er-landm%C3%A6nds-intranet>. Statistik for brug af Facebook er derfor også centralt i et statistikcenter.

Alle de ovenstående datakilder har allerede inden projektets start været vigtige for indsamling af information om vores brugere, men det har været vanskeligt at danne sig et overblik på tværs af platforme, ligesom det ofte har været vanskeligt at sikre, at data er nået ud i organisationen, så der har kunnet handles på dem.

## **2) Gode visuelle præsentationsmuligheder**

Hvis det skal være muligt at danne sig et overblik over en statistik uden forudgående introduktion, er det selvfølgelig vigtigt, at det rent visuelt er nemt og enkelt at aflæse præsentationer.

Et stort problem ved de tidligere nævnte datakilder har netop været, at det kun er en meget begrænset skare, der løbende har fulgt udviklingen i data. Og selvom det løbende er blevet formidlet internt, hvordan man tilgår data de forskellige statistikker, så har systemerne alligevel været for komplekse, til at man har kunnet opnå en udbredelse i SEGES' generelt.

Det er selvfølgelig også vigtigt, at det system der vælges, er i stand til på en lettilgængelig måde at tilbyde en palette af præsentationsformer.

## **3) Gode konfigurationsmuligheder (uden udviklerassistance)**

Flertallet af de datakilder der anvendes, bruges i mange forskellige sammenhænge. F.eks. er der tracking med Google Analytics på alle SEGES' hjemmesider.

Derfor skal det være muligt og enkelt at genbruge den samme opsætning på tværs af datakilder. F.eks. skifte fra præsentation af data for én hjemmeside, til at bruge den samme opsætning til at vise data fra en anden. Dette må ikke – eller kun i meget begrænset omfang – kræve udviklerassistance. Dels for at holde ressourceforbruget nede, dels for at sikre, at der ikke bliver rigtig mange trin og mange forskellige personer der skal involveres, før en visning er på plads.

Samtidig skal det også være muligt for en administrator at ændre i de opstillede visninger, f.eks. at skifte fra et søjlediagram til en kurve, at ændre layout på præsentationen m.v.

## **4) Udbredt og afprøvet platform**

For at sikre, at SEGES i videst mulige omfang bygger videre på kendte teknologier – frem for at kaste mange ressourcer efter udvikling af et selvudviklet system – er det selvfølgelig vigtigt, at det system der vælges, er et, der bruges til løsning af tilsvarende problemstillinger i andre virksomheder.

Samtidig skader det selvfølgelig ikke, hvis man også i SEGES på forhånd kender til det valgte system.

## **5) Ressourcemæssigt overskuelig IT-udviklingsopgave**

For at få mest muligt ud af de forhåndenværende ressourcer, har der været lagt vægt på at bruge kendte teknologier og dermed generelt begrænse mængden af ressourcer, der bruges på IT-udvikling.

Efter at have analyseret en række forskellige systemer – her i blandt som nævnt også muligheden for at bruge ElasticSearch / Kibana – blev Microsoft PowerBI valgt som platform. Det var det system, der bedst honorerede de ovenstående krav, bl.a. er PowerBI meget stærkt på den grafiske præsentation og på administrative konfigurationsmuligheder.

## Udvikling

I forbindelse med introduktionen af PowerBI blev det valgt at lade Novicell være ekstern leverandør og samarbejdspartner. Hos Novicell har de stor erfaring med PowerBI og det blev derfor vurderet til at være mest effektivt, at lade dem stå for implementeringsdelen hvad angår de første rapporter. Efter denne leverance var på plads, kunne ansvaret og de videre udviklingsopgaver overgives til interne ressourcer til SEGES.

### **LANDBRUGSINFO I POWERBI:**

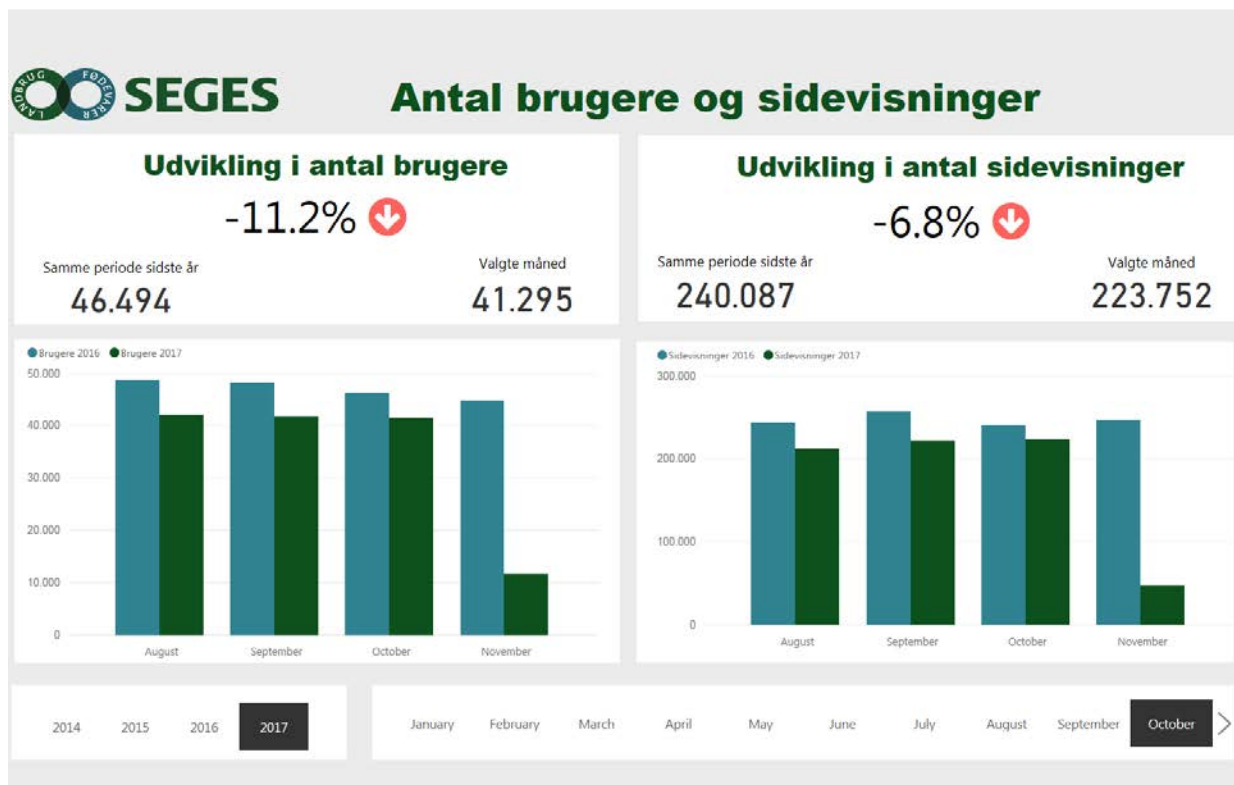
LandbrugsInfo er SEGES' centrale formidlingskanal, og at følge performance for denne hjemmeside, er en meget vigtig del af at sikre, at den viden der ligger i SEGES, også når ud til den enkelte landmand.

Visningen består af 4 separate sider (samt en forside uden data)

1. Antal brugere og sidevisninger
2. Udvikling i antal unikke indloggede. (Separate visninger for hhv. aktive medlemmer af L&F og DLBR-ansatte.)
3. Sidevisninger og kanaler
4. Antal sidevisninger fra LandbrugsInfos nyhedsbrev.

Herunder en gennemgang af indholdet på de enkelte sider:

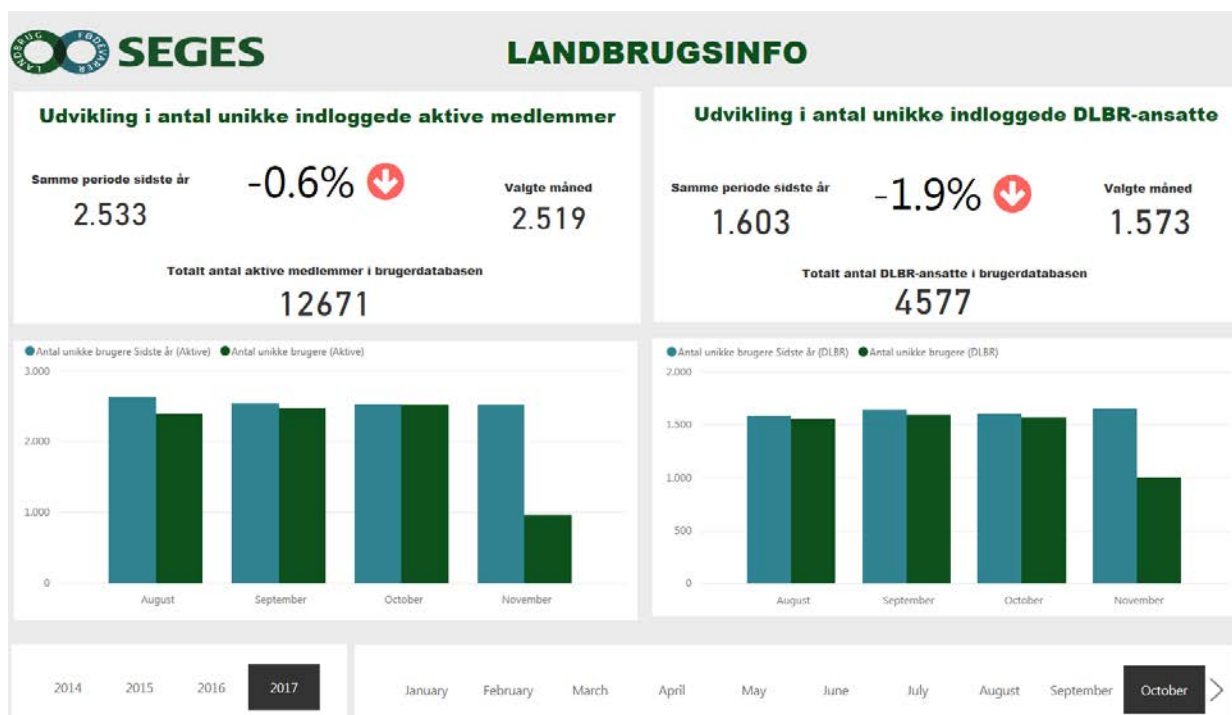
## Antal brugere og sidevisninger:



Denne side giver en generel oversigt over besøgstallene for hjemmesiden. Alle tal, der præsenteres, er taget fra den periode, der er valgt nederst på siden

Øverst i venstre side ser man udviklingen i antallet af brugere i perioden. Dette benchmarkes mod samme periode sidste år. Det samme er tilfældet med "Udvikling i antal sidevisninger" i højre side. Under disse findes en grafisk præsentation af hvordan tallene har udviklet sig i de seneste 4 kalendermåneder, dvs. indeværende måned + de tre forudgående. Som standard står visningen på seneste hele måned, da forandringerne i besøgstal pr. kalendermåned selvfølgelig først kan opgøres endeligt ved månedens udgang.

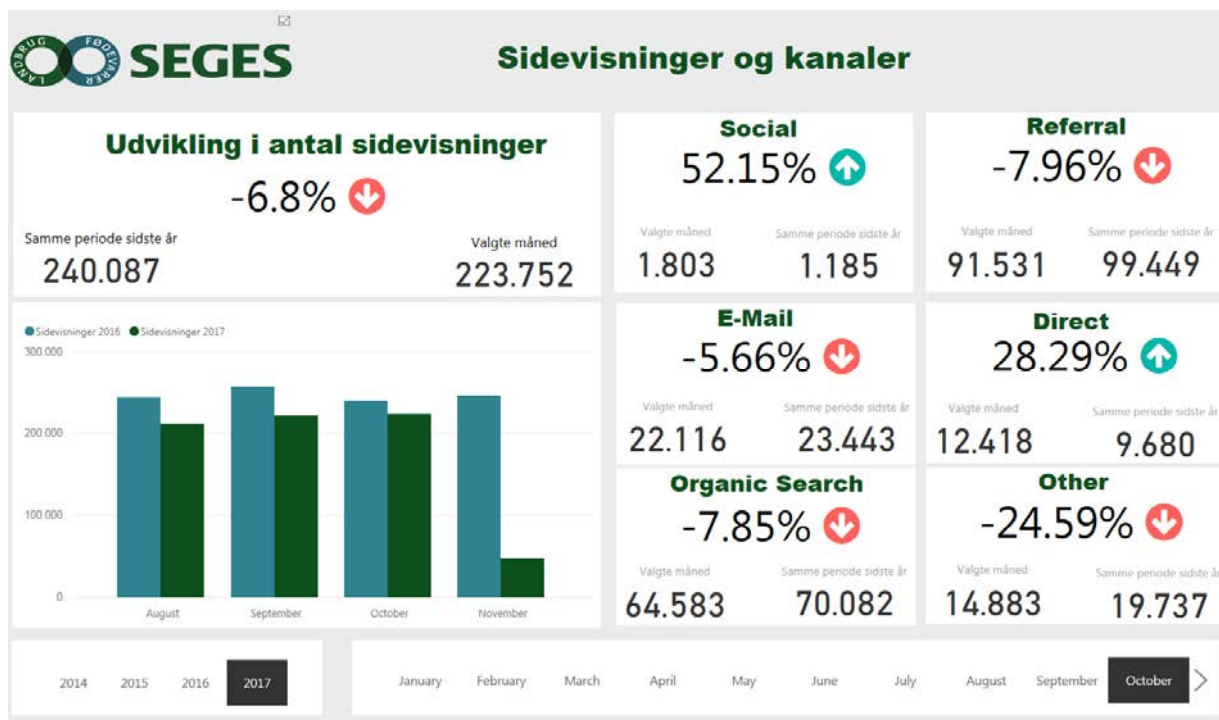
## Udvikling i antal unikke indloggede:



På denne side er der tilføjet en vigtig ny funktionalitet i forhold til tidligere trafikvisninger. Fra SEGES har man i flere år fulgt aktiviteten blandt de indloggede brugere, for at følge op på om bl.a. landmænd og konsulenter brugte hjemmesiden. Men der har manglet indblik i ikke bare hvor stort antallet af brugere har været, men også hvor stor en andel af de potentielle brugere, der har anvendt hjemmesiden. Hvis der f.eks. er et fald i antal sidevisninger til indloggede aktive medlemmer af L&F på 2%, så er det jo ikke nødvendigvis et dårligt tal, hvis antallet af aktive medlemmer samtidig er faldet 5%. Omvendt er en stigning på 10% i antal sidevisninger ikke godt, hvis det totale antal medlemmer af gruppen i samme periode er steget 20%.

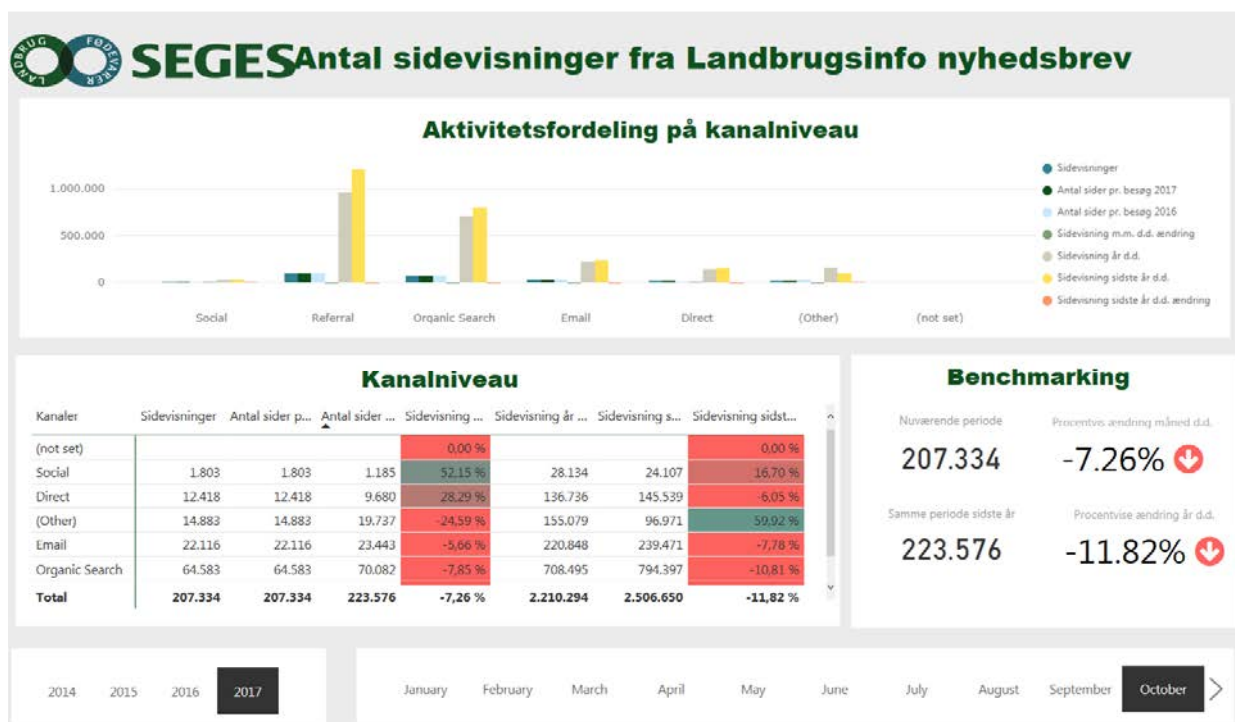
Den ovenstående visning tager derfor udgangspunkt i hvor mange medlemmer, der helt konkret er i brugerdatabase på opgørelsestidspunktet, så man derved får det præcise tal for udviklingen.

## Sidevisninger og kanaler:



På denne side følger man den indkommende trafik til LandbrugsInfo. Den giver overblik over de mange forskellige kanaler brugerne kommer fra – og en benchmark imod hvordan trafikken var fra kilderne i samme periode året før.

## Sidevisninger fra LandbrugsInfos Nyhedsbrev:



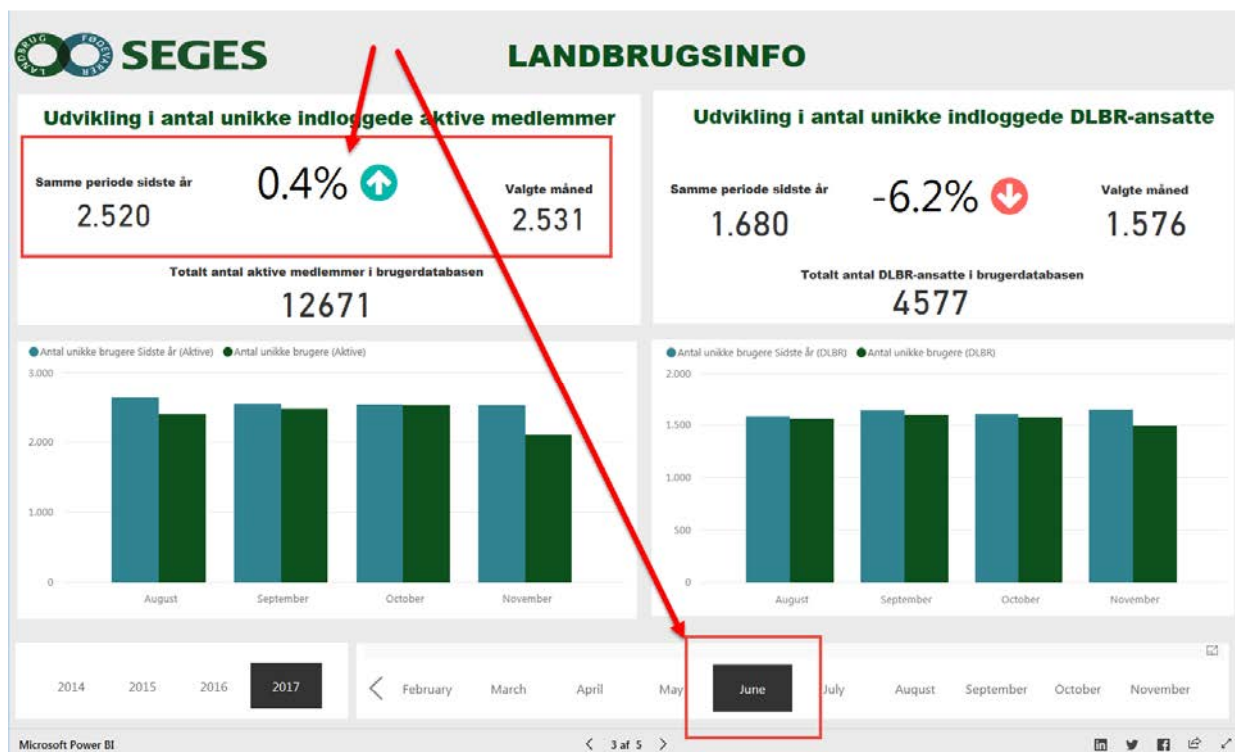
Denne visning sætter spot på aktiviteten fra LandbrugsInfos Nyhedsbreve. Indholdsmæssigt er der ikke arbejdet afgørende med nyhedsbrevene i 2017, da der pågår en omlægning til en ny platform for udsendelse. Denne omlægning afsluttes ultimo 2017, så tallene er primært af interesse for fremtidig benchmarking.

## Interaktive rapporter:

En helt central funktionalitet i PowerBI er, at brugerne ikke blot får vist statistiske sider med data, men at siderne er interaktive. Det giver helt nye muligheder for at analysere på data.

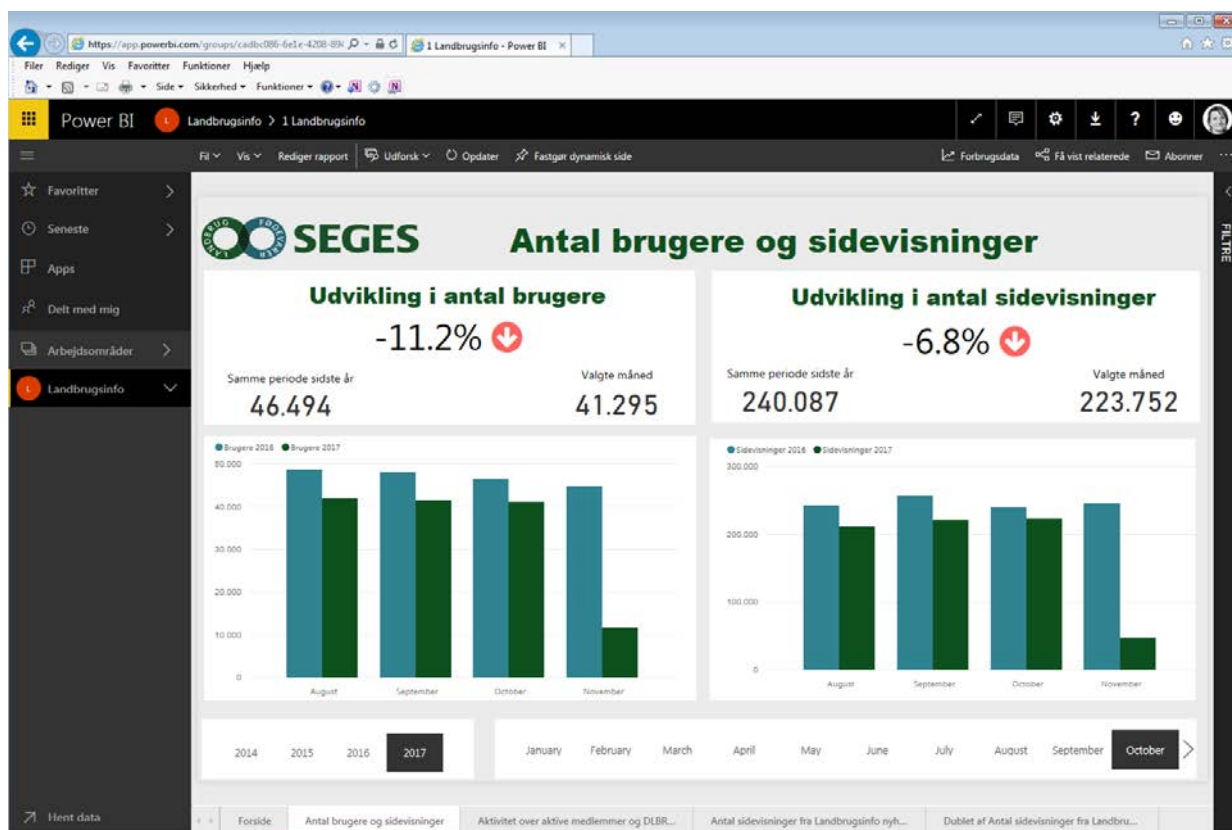
### Eksempel fra oversigten over andelen af indloggede brugere på LandbrugsInfo:

Her er skiftet til juni måned, hvor der ses en positiv udvikling i antal indloggede aktive medlemmer af Landbrugs & Fødevarer, når der benchmarkes mod samme måned i året før:



## Konfiguration med PowerBI:

Den enkelte rapport sættes op og konfigureres vha. en webgrænseflade:



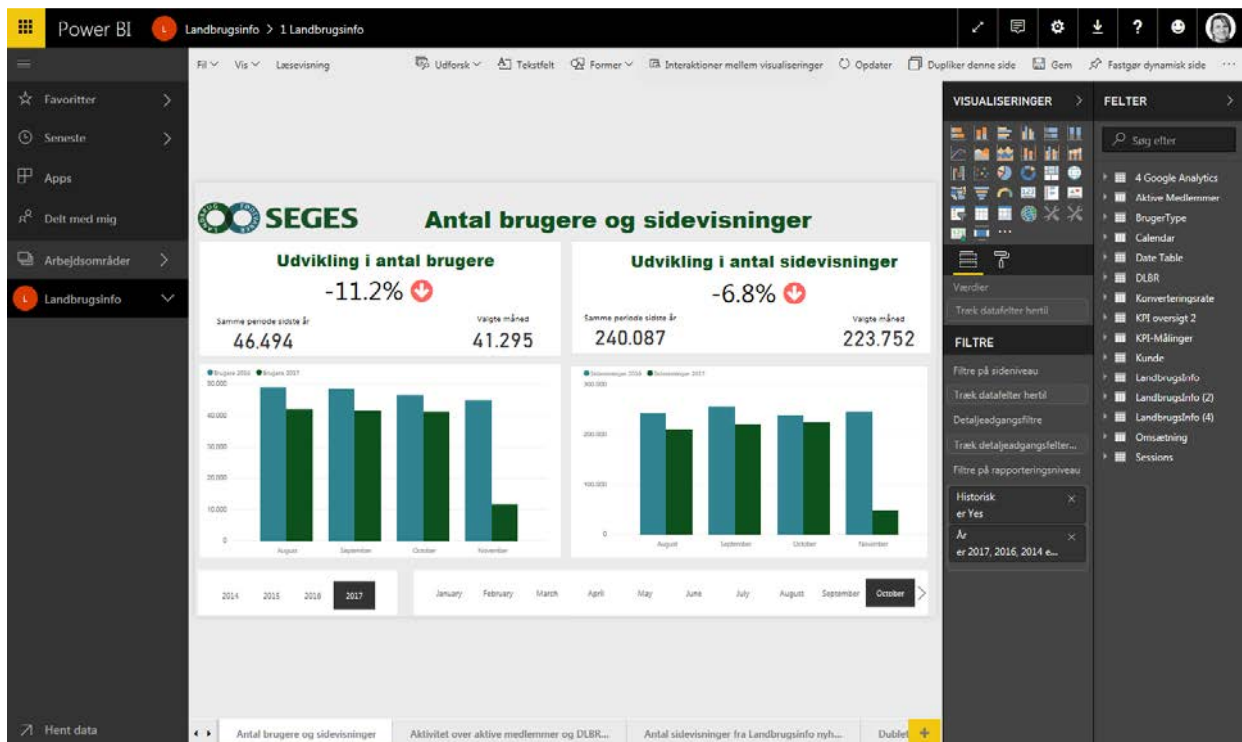
Rettighed til administration af rapporter styres vha. integration til SEGES' egen brugerdatabase. På den måde sikres det, at kun ansatte med de nødvendige rettigheder, har adgang til at opsætte og efterfølgende konfigurere de nødvendige rapporter.

### Redigering af visninger

Administrationsinterfacet til redigering af visningerne giver i vid udstrækning mulighed for personlig konfiguration af rapporter.

Ved redigering af en visning/rapport får man i højre side af administrationsinterfacet mulighed for at redigere hhv. visualiseringer og felter:

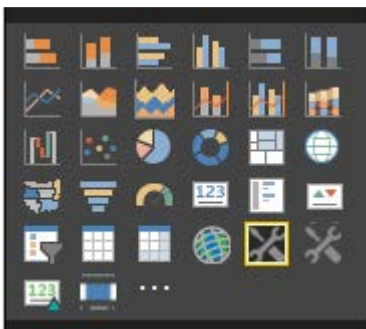




## Redigering af visualiseringer

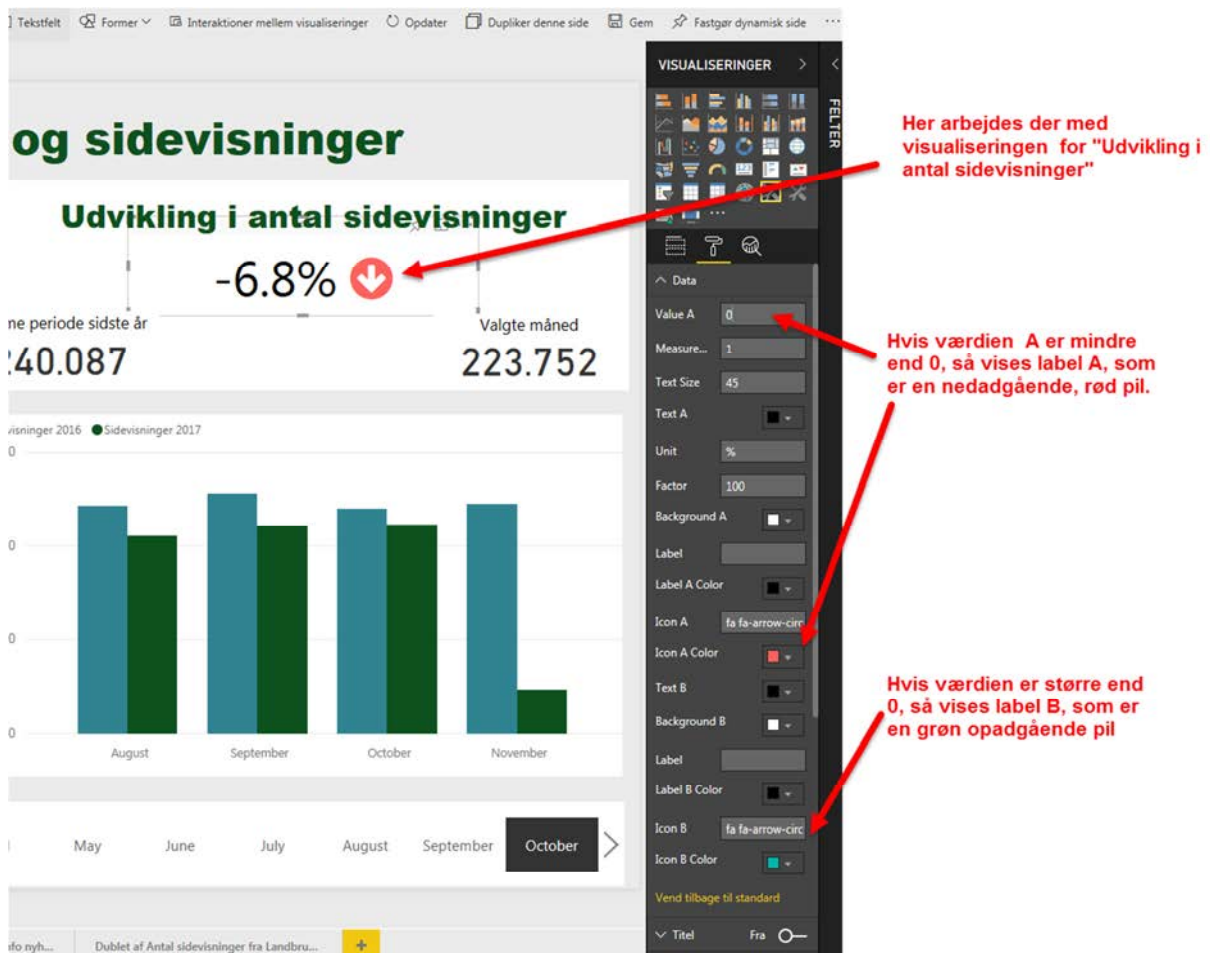
Visualiseringerne redigeres ved, at man udvælger det område, der skal ændres, og derefter konfigurerer visningen.

De grundlæggende præsentationstyper svarer til dem man kender fra f.eks. Excel. Der er således mulighed mellem for at vælge mellem en bred palette af diagrammer, f.eks. lagkagediagrammer, søjlediagrammer, kurver, lister, tabeller osv. Disse vælges på dette panel:



Imidlertid har man også ofte brug for at kunne skræddersy en visning specielt til den enkelte præsentation. F.eks. som med visningen for LandbrugsInfo ovenfor, hvor det er vigtigt at få en tydelig indikator for, om besøgstillere er inde i en positiv eller negativ udvikling. Dette er markeret ved "trend-pile", der viser en rød nedadgående pil ved faldende tal – og omvendt en grøn opadgående pil, når man er inde i en periode med en positiv udvikling.

Dette konfigureres på følgende måde:



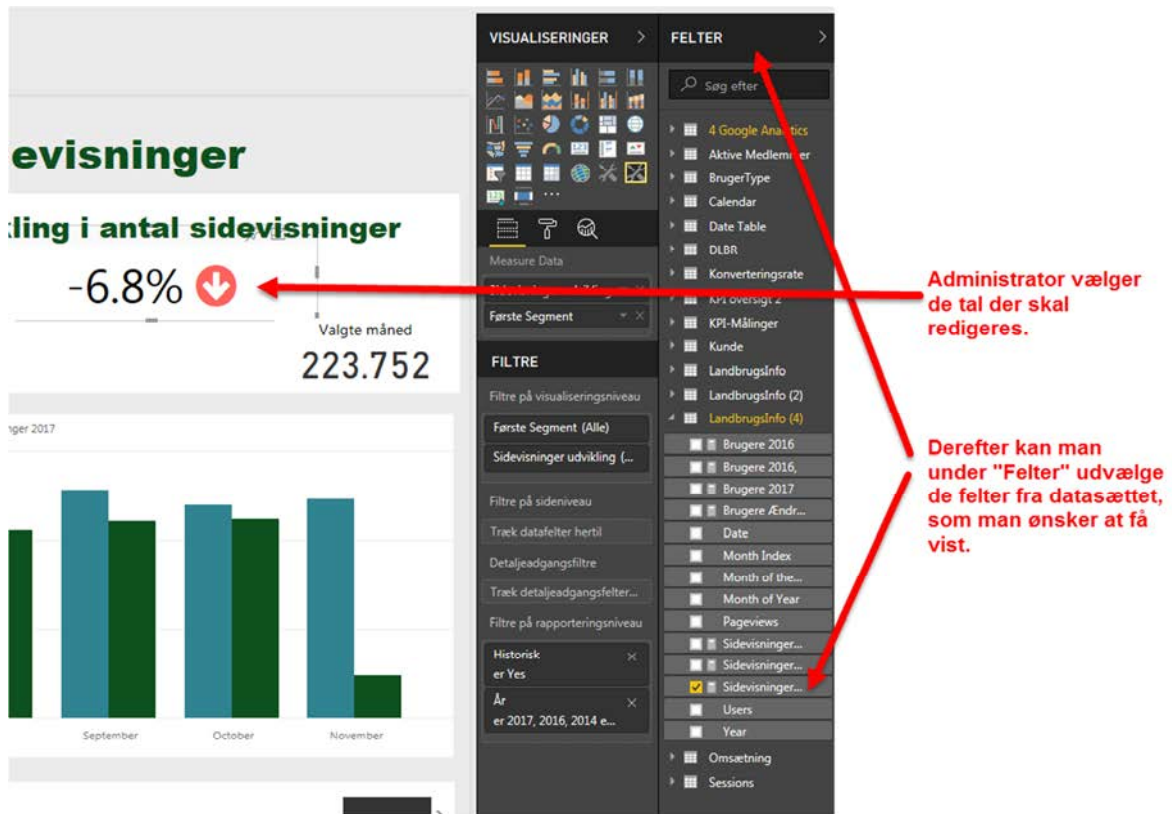
På denne måde er det muligt ikke blot at opsætte og redigere den enkelte præsentation, men også få den tilpasset helt specifikt til de data, der skal vises.

### Redigering af felter

De data, der vises i præsentationerne, kommer fra et eller flere underliggende datasæt. Disse datasæt har en række datafelter der enten kommer direkte fra eksterne kilder som f.eks. Google Analytics eller Navision. Men det er også muligt i PowerBI at kode logik ovenpå de data, der kommer fra de underliggende kilder, sådan at man får de tal frem, som man ønsker at præsentere.

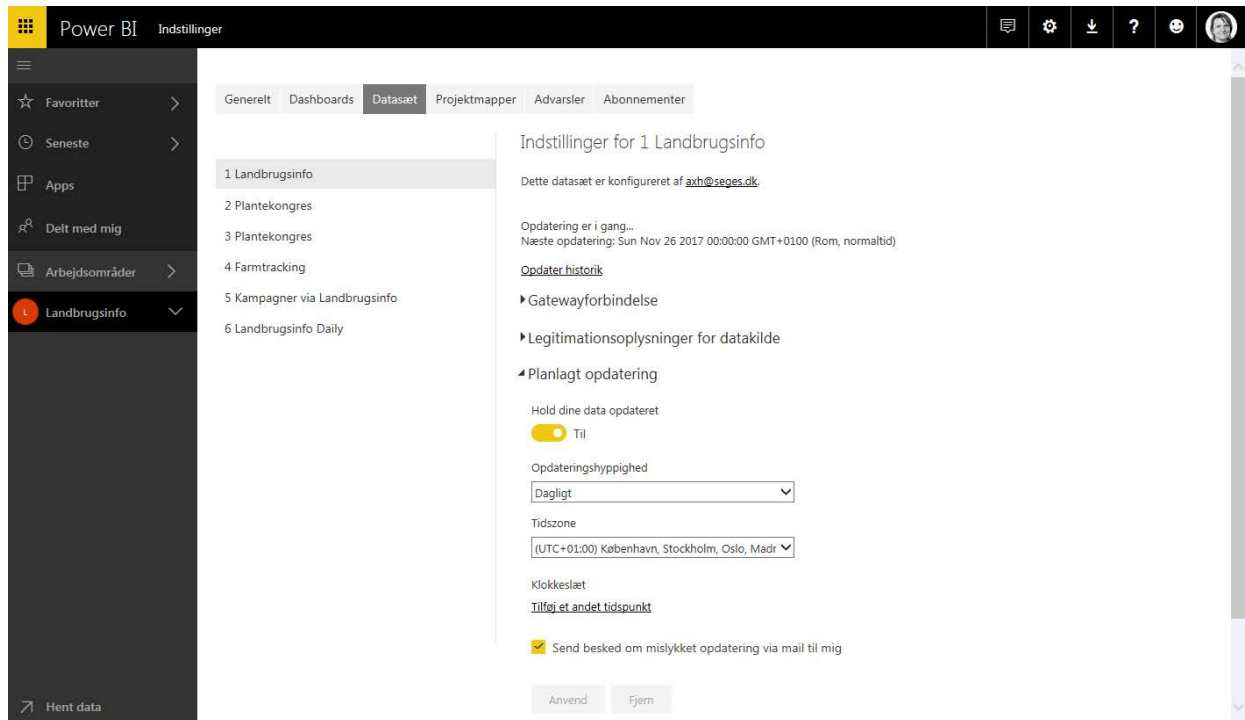
Hvis vi igen ser på eksemplet med "Udvikling i antal sidevisninger", så kommer de underliggende data fra Google Analytics. Her kan man hente antallet af sidevisninger for en bestemt periode. Men det er ikke også muligt at hente den procentvise udvikling i tallet direkte fra Google Analytics. Derfor opstiller man et såkaldt kalkuleret felt, der beregner udviklingen, og dermed gør det muligt for administrator at trække præcis dette tal frem i visningen.

## Eksempel på kalkuleret værdi



## Datakilder

De enkelte datakilder sættes op vha. såkaldte "Gateways" som sørger for integration til de eksterne systemer. Udover at man udvælger hvor data skal hentes fra, er det også muligt at opsætte hvor ofte data skal opdateres. Maksimalt dog 8 gange i døgnet.



## Forbedring af formidling vha. statistikcentret

Statistikcentret og de tilhørende rapporter har været forudsætningen for, at der i løbet af 2017 er blevet skabt bred adgang til data for hele SEGES.

Denne adgang er selvfølgelig nødvendig for, at der kan optimeres på den formidling der foregår.

Statistikcentret bruges på tværs af SEGES' mange platforme, idet man i de fleste situationer blot skifter den underliggende datakilde, men kan genbruge rapporterne.

De datakilder der er arbejdet med i 2017 drejer sig primært om Google Analytics, SEGES' brugerdatabase (Microsoft Active Directory) og nyhedsbrevssystemet Mailchimp.