

Business Analytics:

Datavisualisering

Viktige beslutninger gjøres ofte basert på innsikt i data. Formidling av denne innsikten handler i stor grad om visuell fremstilling ved bruk av grafer og tabeller. Denne femte artikkelen i serien om dataanalyser viser hvordan data bør visualiseres og formidles på en presis og forståelig måte.



Dataingeniør/MBA
Karl Kleive
Senior manager, BDO

Artikkelserie om Business Analytics

I en artikkelserie vil Henning Torgersen og noen av hans kolleger i BDO skrive om hva Business Analytics er og hva det kan brukes til. Her skrives blant annet om datafangst og datahåndtering i revisjon og om hvordan dataanalyser kan brukes både i revisjon og regnskap.



Good communication
doesn't just happen. It is
the result of good design.

- Steven Few

Kommunikasjon med data

Den best egnede visualiseringen av informasjon kan variere, men for å optimalisere tilnærmingen er det viktig å forstå hvordan visuell oppfatning påvirker forståelsen hos brukeren. Denne artikkelen er basert på arbeidet til datavisualiseringsguruven Steven Few¹ og forklarer noen grunnleggende prinsipper for design av effektive visualiseringer av data.

Datadrevne beslutninger

Organisasjoner som ønsker å være datadrevne – altså ta gode beslutninger basert på data – vet at dårlig datakvalitet i verste fall kan føre til feil beslutninger. Organisasjoner bruker mye tid og krefter på å sikre et korrekt data-grunnlag. Data er ikke nyttige i seg selv. De må presenteres på en måte som gir brukeren viktig, meningsfull og nyttig innsikt. Først da har brukeren et godt grunnlag for å ta de riktige beslutningene.

Det er ikke enkelt å finne statistikk om hvor ofte beslutninger er blitt tatt på feil grunnlag som følge av dårlig designede datavisualiseringer. Det er likevel

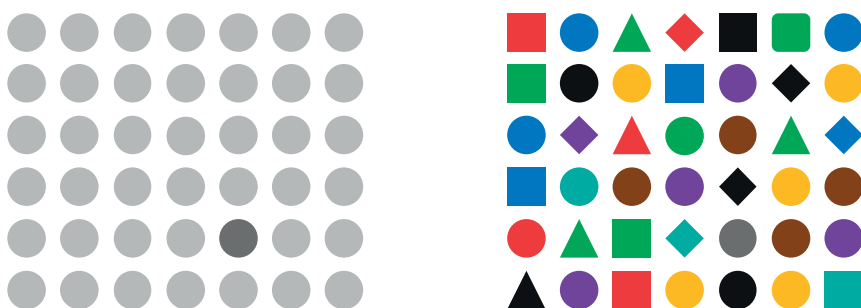
grunn til å anta at mange snubler på målstreken fordi visuell kommunikasjon av informasjon ikke får den oppmerksomheten den fortjener. Det skorter ikke på muligheter. Grafer og tabeller er lett tilgjengelig i de fleste programvarepakker som håndterer data, som for eksempel i Excel, men få av oss har fått opplæring i design av grafer og tabeller.

Gjør det enkelt

Visuell presentasjon av data bør gjøres enkelt. Grafer og tabeller skal være

intuitive og entydige, og forstås ved første blick uten å måtte tolkes.

Det er enkelt å oppdage en ørn i en himmel full av duer, men hvis himmelen er full av mange forskjellige fugler er det vanskeligere å identifisere den ene ørnen. Det samme gjelder for grafer. Det er om å gjøre å begrense antall farger, former og størrelser for å utnytte underbevisstheten slik at brukeren raskt og enkelt forstår den visuelle presentasjonen av data.



Figur 1

Eksempel 1:

Ved å begrense antall farger og former slik som i figuren til venstre, er det enkelt raskt å identifisere objektet som skiller seg ut. Det samme objektet befinner seg i figuren til høyre, men på grunn av mylderet av former og farger er det ikke lenger like enkelt å oppdage den mørkegrå sirkelen.

¹ Steven Few (2012) *Show me the numbers*.



Having the means to create graphs with a computer doesn't guarantee that we'll do it effectively any more than having word processing software makes us great writers.

- Steven Few

Grafer

Grafer setter brukere av informasjon i stand til å se det store bildet. Det er enkelt å sammenligne størrelser og å se trender samtidig som det er lett å avdekke mønstre og avvik.

Valg av graf

Hvilken type graf som best kommuniserer informasjonen, kommer an på dataene, mottaker og hva man ønsker å formidle. Likevel finnes noen rettesnorer basert på hvordan vi mennesker oppfatter former. Ofte er stolper og linjer nok til å dekke de fleste behov.

Kakediagram

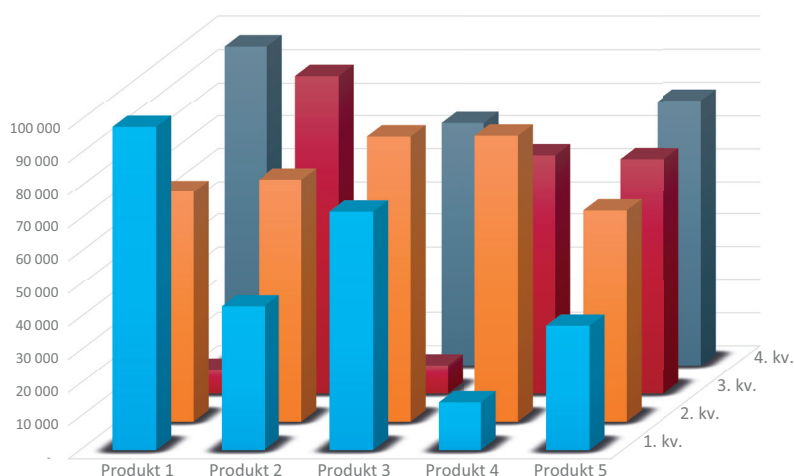
Kakediagram er sjelden det beste alternativet for visualisering av størrelser. Det er fordi det er vanskelig å estimere forholdet mellom to arealer, som vist i eksempel 3.

Det kan være fristende likevel å ty til kakediagrammet ved å legge til prosentandelene til hver kake, men da er det disse tallene som gir innsikten, ikke den visuelle fremstillingen. Å presentere prosentandelene kun i en tabell istedenfor et kakediagram, ville altså ha redusert mengden informasjon hjernen måtte prosessere, og gitt raskere og mer presis innsikt.

Et stolpediagram vil i de aller fleste tilfellene være et mye bedre alternativ ettersom det tar mindre plass i en rapport eller på et dashboard, samtidig som det opprettholder informasjonsverdien.

Stolpediagram

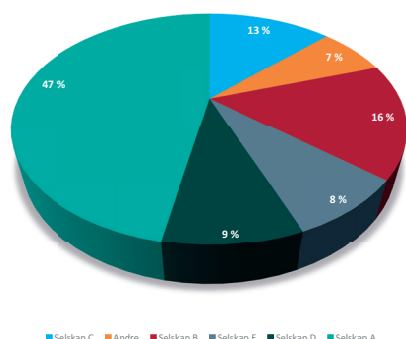
Stolpediagram er godt egnet for å visualisere data som sorteres i ikke-numeriske kategorier. Det er enkelt å kjenne igjen disse kategoriene ettersom de i seg selv ikke lar seg sortere fra minst til



Figur 2

Eksempel 2:

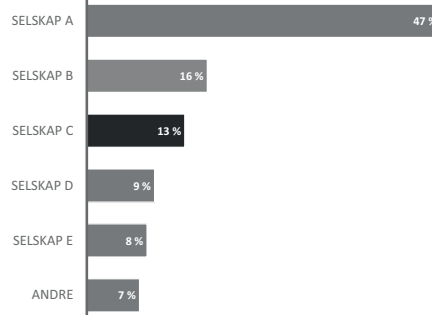
Som figur 2 illustrerer, kan 3D-effekter, farger og skygger gjøre grafer feiende flotte. Det kan være greit for å få et raskt overblikk, men det kompliserer avlesing og reduserer presisjon. Det er ikke enkelt å lese av inntjeningen på Produkt 5 i første kvartal uten å bruke linjal. Og hva var egentlig inntjeningen fra Produkt 2 i fjerde kvartal?



Figur 3

Eksempel 3:

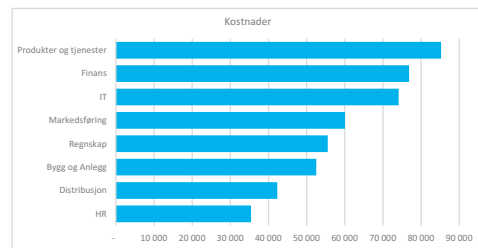
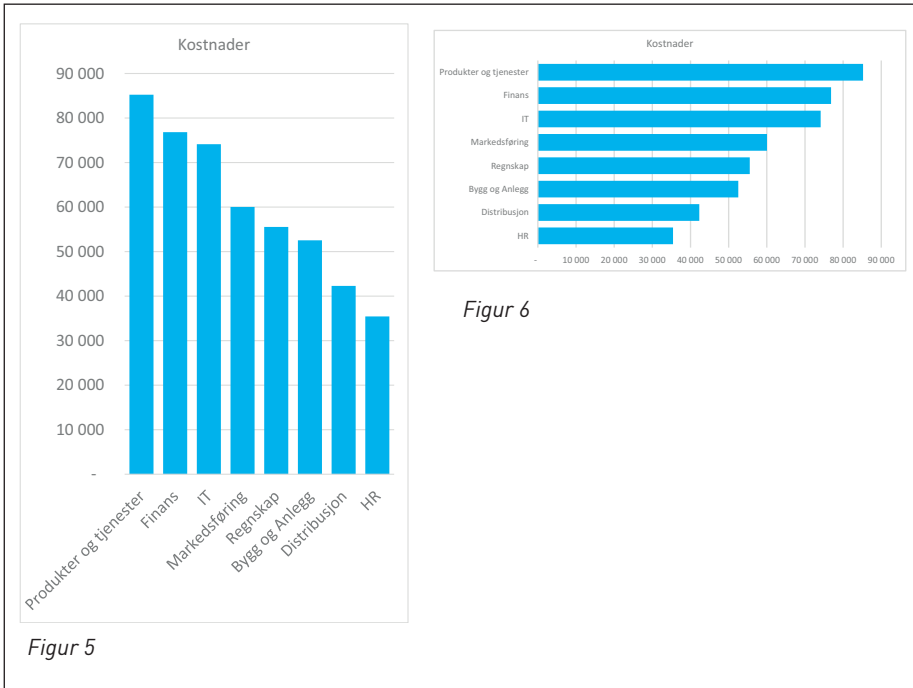
Vi ønsker å vise posisjonen til Selskap C i markedet. I figur 3 er det enkelt å se at Selskap A er størst, men 3D-effekten gjør at kakestykker som er i front oppleves større enn de som er bak, slik som her hvor 9 % ser større ut enn 13 %. Figur 4, derimot, gir oss på en enkel måte oversikt over markedet samt presise markedsandeler. Selv om den mangler flotte farger, kommuniserer dette stolpediagrammet enkelt og elegant all informasjon som brukeren trenger.



Figur 4

For å fremheve det man ønsker å kommunisere, er det lurt å fjerne så mye som mulig av unødvendige eller forstyrrende elementer. 3D-effekter, skygger og bakgrunnsbilder er eksempler på elementer og effekter som faktisk van-

skeliggjør avlesing av diagrammer. Andre eksempler på såkalt «chart junk» er støttelinjer, som kun er et hjelpeverktøy og som med fordel kan være så vidt synlige. En generell anbefaling er å spørre seg selv: «hva kan jeg fjerne?»

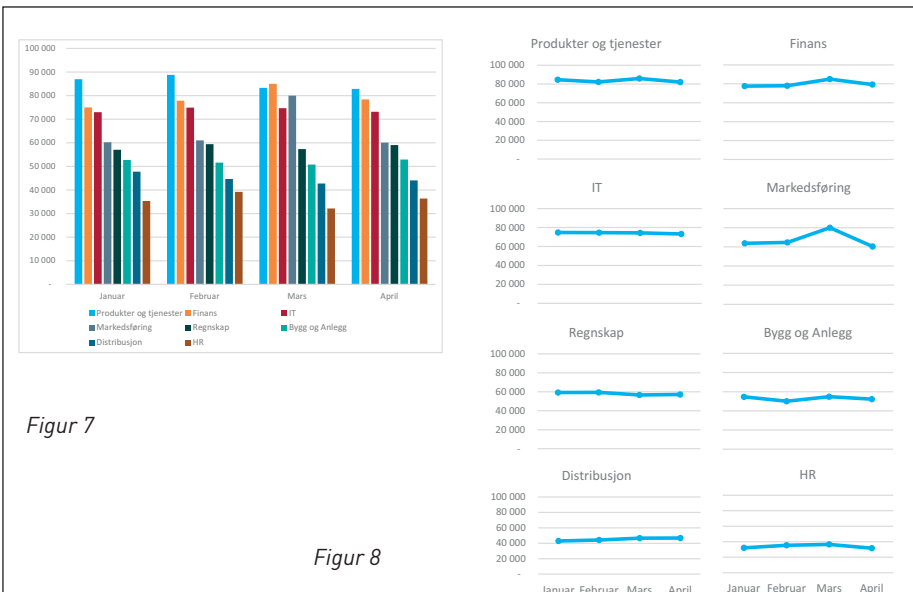


Det er faktisk like lett stående som liggende, bruk derfor gjerne liggende hvis du har lang tekst. Da trenger ikke leseren å bruke unødvendige krefter på å legge hodet på skakke for å prøve å lese den vertikale teksten. Se figur 5 og 6.

Linjediagram

For data med numeriske kategorier er ofte linjediagram det beste, og linjediagram eller punktdiagram blir som oftest brukt for å vise trend over tid.

I likhet med stolpediagrammet, er linjediagram også enkelt og raskt å lese. Dette er fordi oppfatning av posisjoner i 2D slik som linjens helning og retning også er en prosess som skjer i underbevisstheten. Se Figur 8 for eksempler på linjediagrammer.



En beslektet diagramtype er scatter-plott, en type punktdiagram som kan brukes for å vise korrelasjon mellom to verdier og som egner seg godt for raskt å identifisere sammenhenger eller avvik.

Én graf til å forklare alt

Mange variabler kan forlede til forsøk på å vise alt i samme graf. Problemet er at det ofte er vanskelig å lese mange dimensjoner av et datasett i én enkelt graf. Det generelle rådet er derfor å dele opp informasjonen som skal presenteres fremfor å presse alt inn i én graf. Det behøver ikke ta så mye mer plass, og kommuniserer på en mye enklere måte informasjonen som ligger i dataene.

Fargevalg

Fargevalg er viktigere enn mange tror og et område det syndes på, men det finnes noen gode kjøreregler.

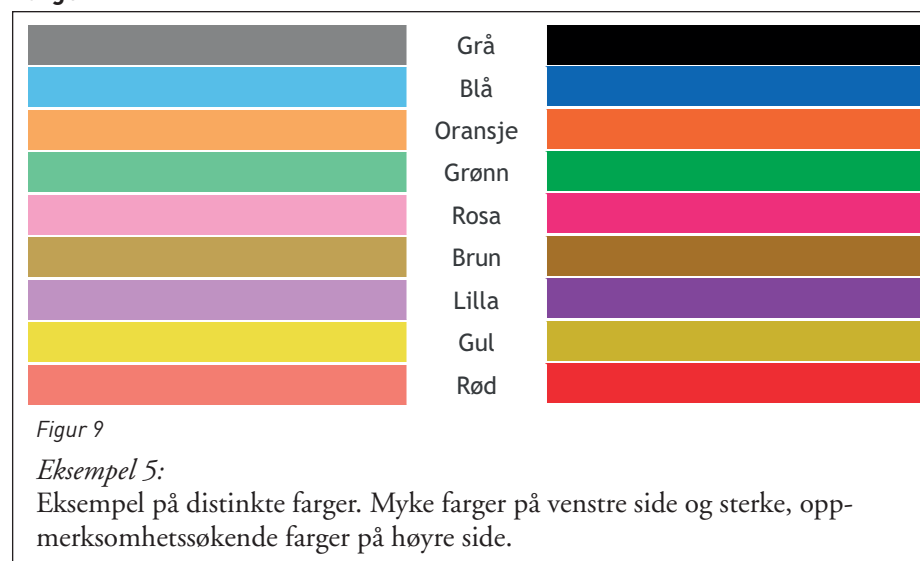
Fargene i en graf bør være distinkte slik at det er enkelt å skille elementene fra hverandre. Videre bør fargene være duse ettersom det er mer naturlig for øyet. Det gjør det også lettere å bruke sterke farger om man ønsker å fremheve spesielle elementer. Er alle farger fremtredende eller sterke, oppfattes dette som at alt skriker etter oppmerksomhet, og resultatet blir at alt drukner.

størst. Avdelinger i en organisasjon er et eksempel på en slik kategori med ikke-numeriske verdier.

Det som gjør stolpediagram så bra, er at det går enkelt og raskt å lese av leng-

der. Fordi vår oppfatning av lengder prosesseres i vår underbevissthet, bruker hjernen minimalt med kapasitet for å sammenligne dem.

Farger



Figur 9

Eksempel 5:

Eksempel på distinkte farger. Myke farger på venstre side og sterke, oppmerksomhetssøkende farger på høyre side.

Farger som har en klar assosiasjon til noe, bør brukes med omhu. Rød assosieres ofte med fare, forbud eller negative tall, mens grønn ofte assosieres med noe positivt.

Er man bundet av organisasjonens fargepalett, er min anbefaling å utfordre tvungen bruk av fargepaletten til datavisualisering. En palett av skrikende farger bør ikke få lov til å komplisere brukerens forståelse. En løsning kan være å ta utgangspunkt i organisasjonens palett og lage varianter av denne med mer eller mindre tint (fargen blandet med mer svart eller mer hvit).

Fargeteori generelt, og i datavisualisering spesielt, er egne fagfelt. Det finnes mange artikler og gratis verktøy for å utforske dette viktige området innen datavisualisering der man kan prøve seg frem og for eksempel se hvordan ens fargepalett vil se ut for en som er fargeblind.

Tabeller

	Q1	Q2	Q3	Q4	Total
Produkt 1	98 130	69 907	7 173	96 889	272 099
Produkt 2	43 346	73 669	96 653	36 756	250 424
Produkt 3	72 223	86 695	8 494	73 706	241 117
Produkt 4	14 377	86 940	72 490	10 112	183 918
Produkt 5	37 526	64 046	71 310	80 353	253 235
	265 602	381 256	256 120	297 816	1 200 793

Figur 10

Ikke-datablekk

	Q1	Q2	Q3	Q4	Total
Produkt 1	98 130	69 907	7 173	96 889	272 099
Produkt 2	43 346	73 669	96 653	36 756	250 424
Produkt 3	72 223	86 695	8 494	73 706	241 117
Produkt 4	14 377	86 940	72 490	10 112	183 918
Produkt 5	37 526	64 046	71 310	80 353	253 235
	265 602	381 256	256 120	297 816	1 200 793

Figur 11

Tabeller

I motsetning til grafer er det vanskelig å sammenligne størrelser, lese trender og oppdage mønstre eller avvik fra en tabell. Det kan likevel være enklere å fremstille informasjon i en tabell selv om grafen ofte er det første man tyr til. Man kan ofte få mer informasjon inn i en tabell, samtidig som tabeller gjør det enkelt å slå opp presise verdier. Figur 10 viser de samme dataene som grafen i eksempel 2 over, men i tabellformat. Det er nå enkelt å finne den eksakte inntjeningen fra Produkt 2 i fjerde kvartal, og vi trenger ikke linjal for å lese av inntjening.

Ikke-datablekk

Ikke-datablekk er kanskje et nytt begrep for noen, men det er ikke avansert. I likhet med chart junk nevnt i innledningen, handler det om å redusere det totale visuelle inntrykket som hjernen må prosessere.

Hvis vi ser for oss at en tabell skal skrives ut med en gammeldags blekkprinter, er ikke-datablekk alt det blekket som brukes på å trykke elementer som *ikke* har med selve dataene å gjøre. Tabellen i Figur 11 viser den samme informasjonen som tabellen i Figur 10, men med mye visuelt fokus på elementer som ikke presenterer data. Det bør være et mål å redusere ikke-datablekk så mye som mulig for å fjerne det unødvendige og fremheve det viktige.

Gjør det enkelt

Etttersom data stadig blir mer tilgjengelige, benyttes de i økende grad til å ta gode beslutninger. Derfor har data ofte stor verdi, og det er viktig at de er korrekte.

Skulle man derimot ikke evne å kommunisere innsikten som dataene representerer på en god måte, er de til liten nytte og verdi.

For å vite mer, er boken «Show me the numbers» av Steven Few et godt sted å begynne for forskningsbasert informasjon om datavisualisering.

Ikke la dårlige grafer, diagrammer og tabeller føre til dårlige beslutninger. Gjør det enkelt.