

1. DATAVÆRDKÆDEN

Nøglen til at forstå, hvordan man arbejder med data, er dataflyten eller **dataværdikæde**. Det ses generelt som sekventielt. Nogle mennesker foretrækker at bruge en livscyklusvisualisering, dette har nogle fordele. I begge tilfælde skal du vedtage et holistisk syn: selv når du bevæger dig trin for trin langs strømmen, har du altid brug for en klar idé om, hvor du vil slutte, inden du starter.

2. DATAINFORMATION: KILDER TIL DATA

Data kan komme fra interne eller eksterne kilder. Hvad er fordelene ved hver datakilde?

- Interne datakilder, hvor data oprettes af virksomheden, er typisk lettere at indsamle og kan være mere relevante for virksomhedens egne formål og indsigt. **1 De er gratis for virksomheden. kan starte en række big data-initiativer uden nogensinde at se ud over deres egne vægge**
- Eksterne data er de, der erhverves uden for virksomheden.
 - o Tredjepartsdata - kan være specifikke for dit marked eller kunder, men vanskelige at opnå og / eller dyrt.
 - o Åbne data - dvs. regeringsstatistikker, data relateret til social påvirkning. Kan være af stor værdi men kræver ofte flere investeringer i analyse for at give indsigt af relevans.
 - o Samlet genererede data. Mange flere virksomheder deltager i Frit og Hackathon.

Casestudie - er eksterne data værd at koste?

- 7 små virksomheder i Nordirland selling forarbejdede fødevarer fået adgang til et stort supermarked kædens data på områder som forbruger livsfase og livsstil, markedsorienteret kurv analyse, og bedste effektive butikker for de små virksomheders produkter.
- Ejere og ledere deltog i workshops for at lære, hvordan man henter og analyserer de mest relevante data fra databasen for at besvare spørgsmål som "Hvordan fungerer min kategori?" "Hvad er den mest populære smag af brød?" "Hvilken type forbruger køber et produkt, der ligner mit?"
- En yoghurtproducent lærte, at ældre voksne var et nøglemarked, så da de besøgte supermarkeder for smag i butikken, forsøgte de ikke længere at lokke yngre kunder og fokuserede i stedet på ældre. Taktikken forbedrede begivenhedernes produktivitet.
- Andre virksomheder kunne forestille sig innovationer i lang rækkevidde snarere end blot at se på andre produkter på markedet og forsøge at efterligne dem eller følge retningslinjerne, der blev fastlagt af de store detailkøbere.
- Dataene forstærkede virksomhedernes iboende iværksætterånd. Arbejdspladser blev mere kollegiale: De fleste af ejerlederne delte kortoplysninger med deres firmaer og opfordrede medarbejderne til at blive involveret og tilbyde nye ideer. ²

¹ <http://www.aunalytics.com/understanding-analytics-part-1-top-internal-sources-of-big-data/>

² <https://www.conference-board.org/blog/post.cfm?post=2526>

Ejere af små virksomheder står over for hindringer for at investere i eksterne data. Barrierer inkluderer omkostninger, deres manglende ekspertise eller tro på deres egen markedskendskab.

Når de fik adgang til dataene, var de fleste af virksomhederne hurtige til at formalisere deres tilgang til marketingplanlægning.

I dette eksempel kunne de små virksomheder få adgang til de eksterne data gratis. Hvor meget kunne de have betalt for, at det stadig var umagen værd?

DATA INFORMATION: TYPE DATA

Data præsenteres som et DATA SET. EN **datasæt** er en gruppe af oplysninger, der er relateret til hinanden.

- Datasæt kan være kvantitative eller kvalitative. Forskellen synes skarp, men der kan være mange måder at få dem til at interagere på. Kvalitative data kan være vanskelige at præcist måle og analysere. Dataene kan være i form af beskrivende ord, der kan undersøges for mønstre eller betydning, undertiden ved hjælp af kodning. Kodning giver forskeren mulighed for at kategorisere kvalitative data for at identificere temaer, der svarer til forskningsspørgsmålene og udføre kvantitativ analyse.
- Data kan være diskrete eller kontinuertlige.
- Data kan være:
 - **nominelt:** Nominelle data er kvalitative; der er ingen iboende skala eller værdi knyttet til dataene.
 - **ordenstal:** Data, der har en naturlig orden, som en Likert-skala, hvor 1 betyder "hader det!" og 5 betyder "elsker det!" Bare husk, at forskellen mellem ordinaler (f.eks. Mellem 1 og 2 eller mellem 3 og 4) ikke nødvendigvis er ens. Ofte er afstanden mellem ekstrem lignende og ikke lide større end mellem at føle sig neutral eller lidt positiv eller negativ.
 - **Interval:** Data, der har en naturlig rækkefølge, og afstanden mellem hver værdi er lig, ligesom temperatur.
 - **Forhold:** Data, der har en naturlig rækkefølge, ens afstand mellem værdier og et naturligt nulpunkt. ³

3. DATA OPBEVARING

Det er vigtigt at tænke fremad, når du planlægger din dataindsamling. Tænk på, hvordan du bruger de data, du indsamler, og hvilke beslutninger, du vil tage, baseret på dem. Ønsker du at få oplysninger om en realtidssituation? Vil du forstå kausalitet? Vil du sammenligne to variabler og vælge den bedste?

Lagring på stedet v fjernlagring

- Traditionelt blev data gemt på stedet, men stigende mængde data og faldende omkostninger for udbydere af datalagringstjenester gør fjernlagring i skyen mere attraktiv.
- Det skybaserede lagermarked domineres af Amazon Web Services, Google og Microsoft Azure.
- Alle opbevaringsskyer oprettes ikke ens. Nogle skyer er optimeret til at håndtere arkivering, andre har ydeevne og stabilitet til at fungere som bagenden for et primært datalagringssystem, og stadig andre er ikke værd risikoen til noget formål. Skyen med den laveste pris kan ende med at blive meget dyrere i det lange løb, hvis data går tabt eller er utilgængelige

³ <https://communitymedicine4asses.com/2013/01/13/scales-of-measurement-nominal-ordinal-intervalratio/>

- Mange løsninger giver fleksibiliteten til at vælge, hvor data skal gemmes: på stedet og / eller i skyen.

4

Hvordan vælger du, hvordan du gemmer dine data?

Når det kommer til lagring af data, er der ingen 'one-size-pas-all'-løsning.

Virksomheder skal forstå mængden og typen af data, de har sammen med motivationen bag lagring af informationen.

HVEM har brug for adgang til hvilken type information? Hvor ofte og hvor hurtigt skal vi få adgang til dataene?

- Når du opretter processer, skal du identificere organisationens vigtigste data og prioritere lagringsstyringsressourcer korrekt. e-mail kan være et virksomheds højeste prioritet, men at gemme og arkivere e-mail-data for en bestemt gruppe, siger direktørerne, kan være mere kritisk end andre grupper
- "Spar penge ved kun at bruge din hurtigste opbevaring, som SSD, til data, som du aktivt bruger, og bruge billigere platforme, som skyen, til at gemme dine arkiverings- eller backupdata,"

HVILK type data skal vi gemme?

- Kan vi være nødt til at kombinere forskellige datasæt i fremtiden?

HVORDAN MANGE data har vi, og hvor meget vil de vokse i fremtiden?

- Opbevar ikke overflødige data. Mange virksomheder indfanger og gemmer overflødige data. Dette var tidligere et mindre problem, når disken var dyr, men med kapacitet, der er mere rimelig, er tendensen at gemme alt.

HVOR LANG skal vi gemme dataene?

- En politik til opbevaring af data er en nødvendighed for både intern datastyring og juridisk overholdelse. Nogle af dine data skal opbevares i mange år, mens andre data muligvis kun er nødvendige i flere dage. "

HVORDAN SIKKER skal dataene være? Hvilke lovkrav skal overholdes?

- Har en plan for genopretning af katastrofer. Uanset hvilken metode eller sikkerhedskopimetode du bruger, skal du huske at teste dem, og sørg for at gendanne dine data.

3 DATAANALYSE

For at besvare dine strategiske vejledende spørgsmål og få en dybere forståelse af din virksomhed skal du tage dine rå data og transformere dem ved hjælp af en kombination af grundlæggende aritmetiske og beskrivende statistikker for at afsløre mønstre - både dem, du forventer at se, og dem, der kan være overraskende. 5

Hvad man skal kigge efter: fakta, statistik og tendenser er alle nyttige, men af forskellige grunde.

Bemærk: Når man taler om datavisualisering, kan det ses som DEL AF den analytiske proces og separat som et middel til at præsentere resultater. De er bundet i hinanden, og den ene kan normalt støtte den anden.

Eksempler på salg og rentabilitet. Dette er ofte et typisk resultat af dataanalyse. (Kilde: Forbes)

4 <https://techcloudlink.com/14-things-you-need-to-know-about-data-storage-management/>

5 <https://www.schusterman.org/playbooks/data/making-meaning/find-patterns-stats/>

Det nedenstående salgs- og rentabilitetsdashboard af Ann Jackson er et visuelt overbevisende, rent designet oversigt over de data, der viser ændringerne over tid, de geografiske forskelle, tab for bestemte produktkategorier og opsummerer nøglepræstationsindikatorer som numre. Ville Ann overlevere dette instrumentbræt til hendes interessenter og kalde det gjort? Nej, for nu starter de virkelige diskussioner. Ann kan sidde med sit publikum og bore i yderligere detaljer at undersøge, hvorfor bestemte resultater har komme omkring og identificere muligheder for at forbedre virksomhedens resultater. Yderligere undersøgelse af dataene, udforskning af dem med dit publikum, det er den værdi, du tilføjer som analytiker, ud over at producere instrumentbræt. De giver dig et fremragende grundlag for disse diskussioner, men burde ikke være slutpunktet. ⁶

4. DATA Tolkning og anvendelse

At have data er ikke nok. Det er ikke nok at analysere data. Det er vigtigt at oversætte den indsigt, man får fra data til handlinger, der understøtter forretningsvækst.

Som med enhver god strategi, er det afgørende for organisationer at bevæge sig fra at være dem, der gør data, til at blive datadriven for at vide, hvornår og hvordan man anvender den, hvornår man skal ændre den, og hvordan man måler dens succes.

Cyklus med brug af data er vedvarende - det stopper ikke bare, når indsigterne præsenteres, men kræver dybdegående diskussioner. Analyse af dataene er et udgangspunkt, ikke et slutpunkt.

På hvilket diagram er der en større forskel?

Dataanvendelse stammer fra videnskab, men det er også en kunst. Data er objektive, men de kan præsenteres subjektivt. I betragtning af at datavisualisering er så vigtig for at give mening om store data og til at tage beslutninger om kørsel, er det vigtigt at forstå, hvordan man gør det.

Slide 1: I det første diagram ser det ud til, at der er en stor forskel.

Slide 2: I dette diagram ser vi begge er en forskel mellem 35,5% og 41,5% - den samme forskel. Dette viser det

- diagrammer kan være vildledende
- det er vigtigt at kommunikere data i formater, der er nøjagtige og i rimelig skala.
- Når oplysninger tilføjes i skala, bliver de meget klarere.

Hvilken måde at vise dataene er mest effektiv på?

Du er nødt til at tænke på, hvordan du kommunikerer data til dine formål OG hvordan du kommunikerer dem tydeligt, så din besked forstås.

I dette eksempel er det farvede diagram det nemmeste at forstå. Hvad ændrede sig? Det originale blå skema var "decluttered" - akserne og etiketterne blev fjernet, og det tomme rum blev reduceret. Derefter blev fremtrædende detaljer "fremhævet" - farverne blev tilføjet for at afspejle et kontinuum af svar, og numrene blev tilføjet for at vise relative værdier.

Konklusion:

Hver gang du bruger data til at formidle noget, fortæller du i bund og grund en historie. Det kan være i en kvartalsvis præsentation for dit bestyrelse, i en formel programevalueringsrapport, i en årsrapport til vælgere.

⁶ <https://www.forbes.com/sites/evamurray/2019/03/29/what-is-the-difference-between-data-analysisand-data-visualization/>

Historiefortælling med data er kraftfuld, og det kommer med et etisk imperativ. Data kan let manipuleres for at fortælle en historie, der ikke er **der, eller for at minimere en historie, der er.** ⁷ **Det er vigtigt at starte med dataene og lade dem fortælle historien i stedet for at lave din historie på forhånd og kun vælge de data, der understøtter den.** Hvis du f.eks. Præsenterer et resumé af en programevaluering, der kun fremhæver de succeser og aspekter af din tjeneste, som folk kunne lide, og ikke inkluderer de komponenter, der kom kort og afsløret områder til forbedring, fortæller du ikke den fulde historie.

Dette projekt er finansieret med støtte fra Europa-Kommissionen. Denne publikation [kommunikation] afspejler kun autorisationens synspunkter eller, og Kommissionen kan ikke holdes ansvarlig for nogen brug, der kan gøres af de informationer, der er indeholdt deri.

⁷ <http://www.bdbanalytics.ir/media/1123/storytelling-with-data-cole-nussbaumer-knafic.pdf>