



# Tiedolla johtaminen ja datatalous

Kirjoittanut

Kira Sjöberg, Esa Tiusanen & Jarmo Rajala

GOODIN

## Sisällysluettelo

<b>Miten organisaatiota johdetaan datalla?</b>	<b>3</b>
Johdanto	3
Mitä on data ja miksi se on tärkeää?	4
Tehtävä 1	6
Datan merkitys liiketoiminnassa	7
Tehtävä 2	8
Tehtävä 3	9
Yhteenveto	10
<b>Miten dataa käsitellään ja hyödynnetään?</b>	<b>12</b>
Datan keräämisen perusteet	12
Tehtävä 4	16
Tehtävä 5	16
Tehtävä 6	17
Datan säilytys ja hallinta	17
Tehtävä 7	21
Datan mallintaminen	23
Tehtävä 8	25
Datan analysointi	26
Tehtävä 9	29
Tekoälyyn perustuva analytiikka	30
Tehtävä 10	33
Datan visualisointi	33
Tehtävä 11	36
Tehtävä 12	36
Yhteenveto	37
<b>Datan voima: Raportointi, mittarit ja eettinen tiedolla johtaminen</b>	<b>39</b>
Suorituskykymittarit strategisen johtamisen tukena	39
Tehtävä 13	41
Reaaliaikainen data kilpailuedun lähteenä	42
Tehtävä 14	44
Kuinka johtaa tiedolla eettisesti?	44
Tehtävä 15	46
Yhteenveto	46
<b>Datan tuomat mahdollisuudet liiketoiminnan tulevaisuudessa</b>	<b>47</b>
Datan kaupallistaminen	47
Tehtävä 16	53
Tulevaisuuden liiketoiminta rakentuu datan varaan	54

Tehtävä 17	56
Data ei ole vain numeroita – se on päätöksiä ja vaikutuksia	56

# Miten organisaatiota johdetaan datalla?

Tämä luku johdattaa sinut datatalouden maailmaan. Saat kokonaiskuvan datataloudesta ja tiedolla johtamisesta: mitä ne tarkoittavat, miksi ne ovat tärkeitä ja miten ne vaikuttavat päätöksentekoon, liiketoiminnan kehittämiseen ja kilpailukykyyn.

## Johdanto

Elämme datan aikakautta, jossa organisaatioiden menestys riippuu yhä enemmän datan tehokkaasta hyödyntämisestä. Tiedolla johtaminen on noussut liiketoiminnan keskiöön, tarjoten mahdollisuuksia tehdä parempia päätöksiä, parantaa tehokkuutta ja löytää uusia kasvumahdollisuuksia. Mutta mitä tiedolla johtaminen oikeastaan tarkoittaa? Entä miten datatalous muokkaa liiketoiminnan toimintamalleja ja kilpailukykyä?

Tämä kurssi tarjoaa kokonaisvaltaisen katsauksen datatalouteen ja tiedolla johtamiseen. Opit, miten dataa voidaan hyödyntää strategisessa ja operatiivisessa päätöksenteossa sekä miten datan laatu, eheys ja sopivuus vaikuttavat organisaation menestykseen. Kurssi auttaa sinua ymmärtämään, kuinka eri datalähteitä – niin organisaation sisäistä kuin ulkoista dataa – voidaan yhdistää ja jalostaa vaikuttavan tiedolla johtamisen perustaksi.

Kurssin aikana tutustut keskeisiin tiedolla johtamisen työkaluihin ja menetelmiin, kuten raportointiin ja data-analytiikkaan. Opit, miten dataa voidaan hyödyntää tietoon perustuvien päätösten tukena. Samalla opit, että tiedolla johtaminen ei ole vain teknisten taitojen hallintaa, vaan se vaatii myös ajattelutavan muutosta. Datalähtöinen ajattelutapa auttaa rakentamaan vahvemman perustan, jolla voidaan luoda arvoa organisaatiolle ja edistää kestäväää kasvua.

Ensimmäisessä luvussa johdatamme sinut tiedolla johtamisen ja datatalouden maailmaan. Opit, mitä data on, miksi se on liiketoiminnalle elintärkeää ja miten datatalous sekä tiedolla johtaminen muokkaavat organisaatioiden toimintaa ja kilpailukykyä. Toisessa luvussa syvennymme siihen, miten dataa käsitellään ja hyödynnetään – käsittelemme datan keräämisen perusperiaatteet, säilytyksen ja hallinnan käytännöt sekä analysoinnin menetelmät. Lisäksi tarkastelemme edistynyttä analytiikkaa ja datan visualisoinnin merkitystä selkeiden ja vaikuttavien johtopäätösten tekemisessä.

Kolmannessa luvussa keskitymme tiedolla johtamiseen ja sen vaikutuksiin. Pohdimme, miten tiedolla johtaminen tukee strategista päätöksentekoa, millainen rooli suorituskykymittareilla on

johtamisen välineinä sekä miten reaaliaikainen data voi toimia kilpailuedun lähteenä. Neljännessä luvussa käsittelemme datan kaupallistamista ja liiketoimintamahdollisuuksia sekä kurkistamme tulevaisuuteen: miten data muovaa liiketoimintaa jatkossa ja millaisia uusia mahdollisuuksia se avaa eri toimialoilla.

Tervetuloa matkalle kohti tiedolla johdettua tulevaisuutta!

## Mitä on data ja miksi se on tärkeää?

Data on nykyaikaisen liiketoiminnan ydin. Se mahdollistaa asiakasymmärryksen syventämisen, tuotteiden ja palveluiden parantamisen sekä toimintojen tehostamisen. Suomessa, yhtenä Euroopan digitalisoituneimmista yhteiskunnista, datan hyödyntäminen on paitsi välttämätöntä, myös merkittävä kilpailuetu. Riippumatta organisaation koosta tai toimialasta, datan tehokas käyttö voi määrittää sen tulevaisuuden ja menestyksen.

Dataa voidaan verrata raaka-aineisiin, kuten öljyyn: raaka öljy itsessään ei ole arvokasta ennen jalostamista, aivan kuten data vaatii analysointia ja tulkintaa tuottaakseen arvoa. Datan todellinen hyöty syntyy, kun sitä käsitellään, yhdistetään ja hyödynnetään päätöksenteon ja toiminnan kehittämisessä.

Dataa syntyy jatkuvasti arjessamme: verkkosivukäynneistä, sosiaalisen median vuorovaikutuksista, älylaitteiden mittauksista ja monista muista lähteistä. Organisaatiot keräävät tätä tietoa ja hyödyntävät sitä toimintansa kehittämiseen, palveluidensa parantamiseen ja uusien liiketoimintamahdollisuuksien luomiseen.

### Käsite

**Data** viittaa mihin tahansa tietoon, joka on kerätty, tallennettu tai käytettävissä analysointia varten. Se voi olla lukumääräistä (kvantitatiivista) tai kuvailevaa (kvalitatiivista) ja se esitetään usein numeroina, teksteinä, kuvina, videoina tai äänenä.

**Datasetti** tarkoittaa tietoa-aineistoa, joka koostuu järjestetystä datasta. Datasetti on usein rakenteellinen kokonaisuus, joka sisältää tietoja, joita voidaan analysoida, jalostaa tai käyttää erilaisissa sovelluksissa ja päätöksenteossa.

Data voidaan jakaa kahteen päätyyppiin: **rakenteelliseen** ja **rakenteettomaan** dataan. Näiden erottelu vaikuttaa siihen, miten dataa voidaan käsitellä, analysoida ja hyödyntää.

## Rakenteellinen data

Rakenteellinen data on järjestettyä ja selkeästi organisoitua tietoa, joka on usein tallennettu taulukoihin tai tietokantoihin. Se soveltuu erityisen hyvin tilastolliseen analyysiin ja automatisoituun käsittelyyn.

Rakenteellinen data on järjestettyä ja selkeästi organisoitua tietoa, joka on usein tallennettu taulukoihin tai tietokantoihin. Se soveltuu erityisen hyvin tilastolliseen analyysiin ja automatisoituun käsittelyyn.

Esimerkkejä rakenteellisesta datasta:

- **Aikasarjadata:** Ajallisesti järjestettyä ja usein tietokantoihin tallennettua dataa. Esimerkiksi myyntiluvut kuukausittain tai säätilan muutokset ajan myötä ovat aikasarjadataa.
- **Operatiivinen data:** Logistiikkaan liittyvät tiedot, kuten toimitusten aikataulut ja varastotasot. Maksutapahtumat, kuten ostot ja palautukset verkkokaupassa ovat esimerkki operatiivisesta datasta.
- **Verkkoanalytiikkadata:** Käyttäjiliikenteeseen liittyvät luvut ja trendit, jotka ovat selkeästi järjestettävissä.

Rakenteellinen data antaa nopeasti selkeitä vastauksia: esimerkiksi toimitusaikoja analysoimalla yritys voi tunnistaa epäkohtia toimitusketjussaan ja parantaa sitä.

## Rakenteeton data

Rakenteeton data on vapaamuotoista ja sen käsittely vaatii enemmän työtä. Sitä ei voi helposti järjestää riveihin ja sarakkeisiin, mutta se voi sisältää merkittävää liiketoiminnallista arvoa.

Esimerkkejä rakenteettomasta datasta:

- **Tekstidata:** sähköpostit, asiakaspalautteet ja sosiaalisen median viestit
- **Kuvadata:** valokuvat tuotteista, dokumenteista tai ympäristöistä
- **Biometrinen data:** kasvonpiirteet ja sormenjäljet, jotka tallennetaan usein kuvia tai äänitiedostoja analysoimalla
- **Emotionaalinen data:** äänensävyistä tai ilmeistä analysoitavat tunnetilat

Rakenteettoman datan avulla voidaan esimerkiksi ymmärtää asiakkaiden tuntemuksia ja odotuksia tuotteista analysoimalla asiakaspalautteita tai sosiaalisen median keskusteluja.

## **Esimerkki**

### **Miten Maija ratkaisi toimitusketjun ongelmat datan avulla**

Maija pyörittää ekologisten vaatteiden verkkokauppaa. Hän on saanut asiakkailta palautetta viivästyneistä toimituksista, mikä on heikentänyt asiakastytyvyyttä. Hän päätti hyödyntää sekä rakenteellista että rakenteetonta dataa ongelman selvittämiseksi.

#### **Rakenteellinen data:**

- Toimitusaikataulut: Maija analysoi toimitusten päivämäärät ja kellonajat sekä viivästystilastot.
- Varaston täyttöasteet: Hän tarkasteli, vastaavatko varastotasot sesonkikysyntää.
- Myyntidata: Historialliset myyntiluvut paljastivat kysynnän vaihtelut.

Näiden tietojen avulla Maija huomasi, että yksi alihankkija oli toistuvasti myöhässä, ja että varaston täydennykset eivät vastanneet kysyntähuippuja.

#### **Rakenteeton data:**

- Asiakaspalautteet: Asiakkaat valittivat esimerkiksi viivästyneistä toimituksista ja epäluottamuksesta toimitusaika-arvioihin.
- Sosiaalisen median keskustelut: Kommenttianalyysi vahvisti viivästysten olevan toistuva ongelma.

Tämän tiedon perusteella Maija neuvotteli tiukemmat toimitusehdot alihankkijan kanssa ja otti käyttöön ennakoivan varastonhallintajärjestelmän, joka hyödyntää historiallista myyntidataa. Hän myös lisäsi verkkokauppaansa selkeämmät toimitusaika-arviot ja viestintäkäytännöt viivästystilanteissa.

Näiden muutosten ansiosta Majan toimitukset nopeutuivat, asiakastytyvyys parani ja palautusten määrä väheni. Samalla hän pystyi optimoimaan toimituskustannuksiaan ja keskittymään liiketoiminnan kasvattamiseen.

Data on perusta niin tiedolla johtamiselle kuin datataloudelle. Kun organisaatio ymmärtää datan eri muodot ja sen hyödyntämismahdollisuudet, se voi siirtyä optimoimaan toimintaansa ja kehittämään uusia liiketoimintamalleja. Tämä ymmärrys on keskeinen seuraavassa osiossa, jossa tarkastelemme datan ja tiedolla johtamisen välistä yhteyttä.

## Tehtävä 1

Valitse oikea vastaus.

### **Miksi dataa verrataan raaka-aineisiin, kuten öljyyn?**

- a) Data on fyysinen resurssi, kuten öljy ja mineraalit.
- b) Data on arvokasta vasta, kun sitä jalostetaan ja hyödynnetään.
- c) Dataa voidaan varastoida rajattomasti kuten öljyä.
- d) Data on uusiutumaton luonnonvara.

*Oikea vastaus: b) Datasta tulee arvokasta vasta, kun se jalostetaan päätöksenteon ja toiminnan kehittämisen tueksi, aivan kuten öljy jalostetaan polttoaineeksi.*

### **Mikä seuraavista on esimerkki rakenteellisesta datasta?**

- a) Sosiaalisen median viestit
- b) Ääninauhitteet asiakaspalvelupuheluista
- c) Valokuvat tuotteista
- d) Myyntiluvut kuukausittain

*Oikea vastaus: d) Myyntiluvut kuukausittain ovat selkeästi järjestettyä rakenteellista dataa, joka voidaan tallentaa ja analysoida helposti.*

## Datan merkitys liiketoiminnassa

Datatalous ja tiedolla johtaminen ovat kaksi toisiaan täydentävää konseptia, jotka ovat muuttaneet organisaatioiden toimintatapoja. Siinä missä datatalous korostaa datan roolia taloudellisen arvon luonnissa, tiedolla johtaminen keskittyy siihen, miten dataa hyödynnetään strategisessa päätöksenteossa ja operatiivisessa toiminnassa. Yhdessä ne tarjoavat organisaatioille työkalut menestyä kilpailuilla markkinoilla ja vastata muuttuvan liiketoimintaympäristön vaatimuksiin.

Datatalous tarkoittaa datan hyödyntämistä taloudellisen arvon luomisessa. Data nähdään tässä yhteydessä resurssina, joka voi synnyttää innovaatioita, uusia liiketoimintamalleja ja merkittäviä parannuksia olemassa oleviin prosesseihin. Esimerkkejä datatalouden käytännöistä ovat:

- uusien tuotteiden ja palveluiden kehittäminen asiakasdatan perusteella

- alustatalouden sovellukset, jotka usein hyödyntävät ennakoivan analytiikan ennustemalleja optimoidakseen asiakaskokemusta
- datan kaupallistaminen, esimerkiksi myymällä dataa tai sen analysoinnin lopputuotoksia muille toimijoille.

Tiedolla johtaminen tuo päätöksentekoon objektiivisuutta, joka vähentää inhimillisten ennakkoluulojen ja "mututuntuman" vaikutusta. Kun päätökset perustuvat konkreettiseen, mitattavaan tietoon, ne ovat paitsi läpinäkyvämpiä, myös paremmin perusteltuja. Tällöin organisaatio voi paremmin hallita riskejä ja tarttua oikeisiin mahdollisuuksiin.

Datan hyödyntäminen tiedolla johtamisessa tuo useita konkreettisia etuja, kuten:

- Pullonkaulojen tunnistaminen ja poistaminen: Prosessidatan analysointi auttaa löytämään tehottomuuksia ja virtaviivaistamaan toimintoja.
- Resurssien älykkäämpi kohdentaminen: Datan avulla voidaan priorisoida tehtäviä, allokoita resursseja ja optimoida kustannuksia.
- Parempi asiakaskokemus: Asiakasdata mahdollistaa palveluiden ja tuotteiden jatkuvan kehittämisen asiakkaiden tarpeita vastaaviksi.

## Käsite

### Tiedolla johtaminen vs. tietojohdaminen – mitä eroa niillä on?

#### Tiedolla johtaminen (englanniksi *data-driven leadership*)

Tiedolla johtaminen keskittyy päätöksenteon tukemiseen hyödyntämällä dataa ja analytiikkaa. Sen tavoitteena on integroida tieto organisaation strategiaan ja operatiiviseen toimintaan. Esimerkki tästä on analytiikkaan perustuva ennakoiva päätöksenteko, jossa dataa käytetään arvioimaan tulevia kehityssuuntia ja tekemään parempia päätöksiä.

#### Tietojohdaminen (englanniksi *knowledge management*)

Tietojohdaminen kattaa organisaation tiedonhallinnan, osaamisen ja oppimisen kehittämisen. Se keskittyy siihen, miten tieto syntyy sekä miten sitä jaetaan ja hyödynnetään organisaatiossa. Esimerkkinä tästä on hiljaisen tiedon jakaminen, oppimisprosessien tukeminen ja tietojärjestelmien hallinta, jotka auttavat luomaan jatkuvasti kehittyvän ja oppivan organisaation.

Siirtyminen kohti dataohjautuvaa organisaatiota vaatii kulttuurin muutosta. Tiedolla johtaminen ja datatalous eivät ole pelkästään teknisiä konsepteja, vaan ne edellyttävät johdon sitoutumista, osaamisen kehittämistä ja teknologian hyödyntämistä. Kun organisaatio ymmärtää, miten dataa

voidaan käyttää arvonluontiin ja päätöksenteon tukena, siitä tulee kilpailukykyisempi ja valmiimpi kohtaamaan tulevaisuuden haasteet.

## Tehtävä 2

Valitse oikea vaihtoehto.

### **Mitä hyötyä tiedolla johtamisesta on päätöksenteossa?**

- a) Se perustuu mututuntumaan ja nopeaan reaktioon.
- b) Se tuo objektiivisuutta ja vähentää ennakkoluulojen vaikutusta.
- c) Se mahdollistaa päätösten tekemisen ilman analytiikkaa.
- d) Se lisää asiakaspalautteen määrää.

*Oikea vastaus: b) Tiedolla johtaminen perustuu dataan ja analytiikkaan, mikä tekee päätöksenteosta objektiivisempaa ja perustellumpaa ja yhdistettynä ihmisten kokemuseräiseen tietoon tuottaa lisäarvoa toiminnalle.*

### **Miten datatalous vaikuttaa organisaatioihin?**

- a) Se mahdollistaa uusien liiketoimintamallien kehittämisen.
- b) Se tekee perinteisistä asiakastiedoista hyödyttömiä.
- c) Se vähentää yritysten tarvetta analytiikalle.
- d) Ei mitenkään.

*Oikea vastaus: a) Datatalous tarkoittaa datan hyödyntämistä taloudellisen arvon luomisessa. Se voi synnyttää innovaatioita, uusia liiketoimintamalleja ja parantaa päätöksentekoa kaikilla toimialoilla.*

### **Mitä eroa on tiedolla johtamisella ja tietojohdamisella?**

- a) Tiedolla johtaminen keskittyy datan hyödyntämiseen päätöksenteossa, kun taas tietojohdaminen keskittyy tiedon jakamiseen ja hallintaan.
- b) Tietojohdaminen on teknistä datan käsittelyä, kun taas tiedolla johtaminen on pelkästään ihmisten välistä viestintää.
- c) Tiedolla johtaminen koskee vain suuria yrityksiä, mutta tietojohdaminen sopii kaikille.
- d) Ei ole eroa – molemmat tarkoittavat samaa asiaa.

*Oikea vastaus: a) Tiedolla johtaminen tarkoittaa datan ja analytiikan käyttöä strategiseen ja operatiiviseen päätöksentekoon, kun taas tietojohdaminen liittyy organisaation tiedonhallintaan, oppimiseen ja hiljaisen tiedon jakamiseen.*

### Tehtävä 3

Valitse, onko väittämä oikein vai väärin.

**Tiedolla johtaminen edellyttää vain teknistä osaamista, eikä sillä ole merkitystä liiketoimintastrategialle.**

- a) Oikein
- b) Väärin

*Oikea vastaus: Väärin. Tiedolla johtaminen yhdistää sekä teknisen analytiikan että strategisen ajattelun. Sen avulla organisaatiot voivat tehdä perusteltuja päätöksiä ja kehittää liiketoimintaa.*

**Tiedolla johtaminen on hyödyllistä vain isoille yrityksille, eikä sillä ole merkitystä pienille ja keskisuurille yrityksille.**

- a) Oikein
- b) Väärin

*Oikea vastaus: Väärin. Kaiken kokoiset organisaatiot voivat hyötyä tiedolla johtamisesta. Pienet ja keskisuuret yritykset voivat parantaa tehokkuuttaan ja kilpailukykyään hyödyntämällä dataa päätöksenteossa.*

### Yhteenveto

**Ensimmäisessä luvussa opimme:**

- Data on nykyaikaisen liiketoiminnan ydin, ja sen arvo syntyy vasta, kun sitä analysoidaan ja hyödynnetään päätöksenteon ja toiminnan kehittämisen tukena.
- Rakenteellinen data on järjestettyä ja helposti analysoitavaa, kun taas rakenteeton data on vapaamuotoisempaa ja sen käsittely vaatii enemmän työtä. Rakenteeton data voi kuitenkin tarjota syvällisempää ymmärrystä esimerkiksi asiakaspalautteiden tai sosiaalisen median viestien analysoinnin kautta.

- Datatalous perustuu datan hyödyntämiseen taloudellisen arvon luomisessa, esimerkiksi uusien liiketoimintamallien kehittämisessä, datan kaupallistamisessa ja alustatalouden ennakoivassa päätöksenteossa.
- Tiedolla johtaminen yhdistää datan ja analytiikan organisaation strategiaan ja päätöksentekoon, mikä mahdollistaa objektiivisemmat, perustellummat ja tehokkaammat päätökset.
- Tietojohtaminen puolestaan keskittyy tiedonhallintaan, osaamisen jakamiseen ja organisaation oppimisen kehittämiseen. Tietojohtamisella varmistetaan, että tieto virtaa tehokkaasti organisaation sisällä ja tukee jatkuvaa kehitystä.
- Siirtyminen dataohjautuvaan toimintaan vaatii kulttuurin muutosta, johdon sitoutumista ja teknologian tehokasta hyödyntämistä, mikä parantaa organisaation kilpailukykyä ja valmiutta kohdata tulevaisuuden haasteet.

## Miten dataa käsitellään ja hyödynnetään?

Jotta dataa voidaan hyödyntää tehokkaasti liiketoiminnassa, sen tulee olla laadukasta, turvallisesti säilytettyä ja analysoitavissa oikeilla menetelmillä. Tässä luvussa perehdymme datan keräämiseen, säilyttämiseen, analysointiin ja visualisointiin.

### Datan keräämisen perusteet

Kuten opimme ensimmäisestä luvusta, data on kaiken tiedolla johtamisen ja datatalouden perusta. Kun kerätty data on laadukasta, relevanttia ja ajantasaista, se luo vankan pohjan, jonka varaan organisaation päätöksenteko ja strateginen suunnittelu voidaan rakentaa. Ilman luotettavaa ja laadukasta dataa analyysit, päätökset ja liiketoiminnan kehittäminen jäävät heikoiksi. Mutta mitä datan kerääminen tarkoittaa käytännössä, ja miksi se on niin tärkeää?

Datan kerääminen viittaa prosessiin, jossa tietoa hankitaan eri lähteistä organisaation käyttöön. Tämä tieto voi olla peräisin esimerkiksi asiakkaiden vuorovaikutuksista verkkosivustolla, tuotantolaitteiden sensoreista tai myyntipisteiden tapahtumista. Keräämisen tavoitteena on saada kattava kuva organisaation toiminnasta, asiakkaiden käyttäytymisestä ja ympäristöstä, jotta dataa voidaan hyödyntää päätöksenteon tukena.

### Datan keräämisen lähteet

Käydään seuraavaksi läpi esimerkkejä tyypillisistä datalähteistä:

- **verkkosivustot ja sovellukset** seuraavat käyttäjäliikennettä, klikkauksia ja muuta verkkokäyttäytymistä
- **asiakaskyselyt ja asiakkuudenhallintajärjestelmät (englanniksi CRM Customer Relationship Management)** tallentavat tietoa asiakkaiden mieltymyksistä ja vuorovaikutuksista yrityksen kanssa
- **ERP- ja toiminnanohjausjärjestelmät** kokoavat ja hallinnoivat yrityksen operatiivista dataa, kuten varastonhallintaa, tuotantoa ja resurssien käyttöä
- **tilausjärjestelmät** seuraavat myyntitilauksia, asiakastoimituksia ja hankintaprosesseja
- **markkinadata** on saatavilla julkaistuista raporteista, kuten Tilastokeskuksen analyysistä, kilpailijoiden julkisista taloustiedoista ja verkkosivuista, sosiaalisen median seurannasta, digitaalisista työkaluista (esimerkiksi Google Trends) sekä kuluttajatutkimuksista ja talousmediasta

- **talousdata** on kirjanpidon ja myyntiraporttien tuottamaa tietoa organisaation taloudellisesta tilanteesta
- **sensorit** keräävät reaaliaikaista tietoa laitteiden toiminnasta ja ympäristöstä, kuten lämpötilasta, kosteudesta tai liikkeestä.

Dataa kerätessä tärkein tekijä on datan laatu. Laatu on tiedolla johtamisen ydin, sillä vain luotettava ja tarkka data mahdollistaa oikeiden päätösten tekemisen. Ilman laadukasta dataa yrityksen strategiat voivat perustua virheellisiin tietoihin tai pelkkiin oletuksiin, jos dataa ei hyödynnetä lainkaan.

## Opi lisää

### Laadukkaan datan keräämisen periaatteet

- **Relevanssi:** Kerättävä data on valittava liiketoiminnan tarpeiden mukaan. Mikä tieto todella auttaa saavuttamaan tavoitteet?
- **Tarkkuus:** Dataa ei voida hyödyntää, jos se sisältää virheitä tai epätarkkuuksia. Esimerkiksi väärät asiakastiedot voivat johtaa huonoon asiakaskokemukseen.
- **Ajantasaisuus:** Vanha data voi olla käyttökelvotonta, joten on tärkeää kerätä ja päivittää tietoa jatkuvasti.
- **Eheys:** Datan tulee olla yhtenäistä ja johdonmukaista, jotta sitä voidaan analysoida tehokkaasti.

Datan keräämisen tulee aina olla tarkoituksenmukaista ja noudattaa voimassa olevia lakeja sekä eettisiä periaatteita. Organisaation tulisi kerätä vain sellaista dataa, joka palvelee selkeää liiketoiminnallista tai kehitystavoitetta. Ylimääräisen datan kerääminen ei ainoastaan kuluta resursseja turhaan, vaan se voi myös vaikeuttaa olennaisten tietojen hyödyntämistä.

Luotettava data on kaiken päätöksenteon perusta. Jotta dataa voidaan hyödyntää tehokkaasti, sen alkuperä ja oikeellisuus tulee dokumentoida tarkasti. Tämä tarkoittaa esimerkiksi sitä, että asiakastietojen keräämisen lähteet kirjataan ja niiden luotettavuus tarkistetaan vertaamalla tietoja toiseen luotettavaan järjestelmään tai pyydetään asiakasta itse vahvistamaan tiedot lomakkeen täyttämisen yhteydessä.

## Opi lisää

### Datan anonymisointi

Anonymisointi datassa tarkoittaa, että tiedosta poistetaan tai muokataan yksilöiviä tietoja, kuten nimiä, osoitteita tai henkilötunnuksia, jotta yksittäistä henkilöä ei voida tunnistaa suoraan tai

välillisesti. Esimerkiksi, kun sairaalan potilastietoja käytetään tutkimuksessa, potilaiden nimet ja henkilötunnukset voidaan poistaa tai korvata tunnisteilla.

PII (englanniksi *Personally Identifiable Information*) tarkoittaa henkilötietoja, joista yksittäinen henkilö voidaan tunnistaa. Tällaisia tietoja ovat esimerkiksi nimi, osoite, puhelinnumero, sähköpostiosoite, henkilötunnus ja IP-osoite. Anonymisoinnilla pyritään muuttamaan PII-tiedot siten, että ne eivät enää liity tunnistettavaan henkilöön, mikä suojaa yksityisyyttä ja mahdollistaa datan turvallisen käytön esimerkiksi tutkimustarkoituksissa.

## GDPR ja sen vaikutus datan käyttöön

GDPR, eli EU:n yleinen tietosuoja-asetus, asettaa selkeät vaatimukset sille, miten henkilötietoja saa kerätä, käsitellä ja säilyttää. Asetuksen tarkoituksena on suojata yksilöiden oikeuksia ja lisätä datan käsittelyn läpinäkyvyyttä. Organisaatioille tämä tarkoittaa paitsi sääntöjen noudattamista, myös mahdollisuutta erottautua vastuullisella ja eettisellä tiedonkäsittelyllä.

Datan keräämisen tulee aina perustua tarpeellisuuteen. Tämä tarkoittaa, että yritykset voivat kerätä vain sellaista tietoa, joka palvelee selkeitä liiketoiminnallisia tai kehitystavoitteita. Ylimääräinen data ei vain kuormita järjestelmiä ja kuluta resursseja, vaan se voi myös hämärtää olennaisten tietojen hyödyntämistä. GDPR painottaa myös avoimuutta: yksilöiden on tiedettävä, mitä tietoja heistä kerätään, mihin niitä käytetään ja kuinka kauan niitä säilytetään.

GDPR:n keskeinen vaatimus on suostumuksen hankkiminen. Asiakkaiden suostumus datan keräämiseen on hankittava selkeästi, ja heidän oikeuksiaan (kuten oikeutta tulla unohdetuksi) on kunnioitettava. Lisäksi henkilötietojen suojaaminen on ensisijaisen tärkeää. Tämä tarkoittaa asianmukaisia teknisiä ratkaisuja, kuten tietojen salaamista ja pääsynhallintaa, mutta myös jatkuvaa prosessien tarkastelua.

### Opi lisää

#### GDPR:n vaatimukset datan keräämiselle ja käytännön toteutus

EU:n yleinen tietosuoja-asetus (GDPR) asettaa tiukat säännöt henkilötietojen keräämiselle, käsittelylle ja säilyttämiselle. Tämä edellyttää organisaatioilta vastuullisuutta ja läpinäkyvyyttä datan käsittelyssä. Alla esittelemme GDPR:n keskeiset periaatteet ja käytännön vinkkejä niiden toteuttamiseen.

#### 1. Tarkoituksenmukaisuus

**Periaate:** kerää vain dataa, joka on liiketoiminnan kannalta välttämätöntä.

### **Käytännön toteutus:**

- laadi selkeä lista datatyypeistä, joita tarvitset liiketoiminnallesi (esimerkiksi nimi, sähköpostiosoite, ostohistoria)
- vältä ylimääräisten tietojen, kuten syntymäajan tai henkilötunnuksen, keräämistä ilman selkeää perustetta.

## **2. Avoimuus**

**Periaate:** ilmoita käyttäjille, mitä dataa kerätään, mihin sitä käytetään ja kuinka kauan sitä säilytetään.

### **Käytännön toteutus:**

- luo selkeä tietosuojaseloste, joka on helposti löydettävissä verkkosivustollasi
- esitä kerätyt tiedot yksinkertaisessa muodossa, esimerkiksi "Keräämme sähköpostiosoitteen uutiskirjeen lähettämistä varten".

## **3. Suostumus**

**Periaate:** hanki selkeä ja yksiselitteinen suostumus henkilötietojen keräämiseen.

### **Käytännön toteutus:**

- lisää verkkosivustosi lomakkeisiin erillinen valintaruutu, jonka käyttäjä voi rastittaa antaakseen suostumuksensa
- varmista, että suostumus ei ole oletuksena valittuna.

## **4. Dokumentointi**

**Periaate:** kaikki kerätyt tiedot ja niiden alkuperä tulee dokumentoida, jotta tiedon käsittely on jäljitettävissä.

### **Käytännön toteutus:**

- hyödynnä CRM-järjestelmiä tai muuta asiakashallintajärjestelmää tiedon tallentamiseen ja dokumentointiin
- kirjaa jokaisen tiedonkeruun yhteydessä myös keräyksen ajankohta ja lähde.

## **5. Oikeellisuus**

**Periaate:** varmista, että kerätty data on tarkkaa ja ajankohtaista.

### **Käytännön toteutus:**

- käytä automatisoituja järjestelmiä, jotka ilmoittavat, jos tietoja ei ole päivitetty tietyn ajan sisällä

- tarjoa asiakkaille mahdollisuus tarkistaa ja päivittää omia tietojaan esimerkiksi verkkopalvelun kautta.

## 6. Tietoturva

**Periaate:** suojaa kerätty data teknisin ja organisatorisin ratkaisuin, kuten salauksella ja pääsynhallinnalla.

### Käytännön toteutus:

- ota käyttöön SSL-sertifikaatti verkkosivustollesi varmistaaksesi turvallisen tiedonsiirron
- rajoita henkilötietoihin pääsy vain niille työntekijöille, joilla on siihen perusteltu tarve.

## Tehtävä 4

Valitse kaikki oikeat vaihtoehdot.

### Mitkä ovat laadukkaan datan keräämisen periaatteet?

- a) Relevanssi
- b) Tarkkuus
- c) Kaiken mahdollisen datan kerääminen varmuuden vuoksi
- d) Ajantasaisuus
- e) Eheys

*Oikea vastaus: a) b) d) ja e). Laadukkaan datan keräämisen periaatteisiin kuuluvat relevanssi, tarkkuus, ajantasaisuus ja eheys, jotka kaikki varmistavat, että data on hyödyllistä ja luotettavaa päätöksenteossa. Ylimääräisen datan kerääminen varmuuden vuoksi ei ole tarkoituksenmukaista.*

## Tehtävä 5

Valitse oikea vaihtoehto.

### Miksi datan kerääminen on liiketoiminnalle tärkeää?

- a) Koska se on laissa vaadittu.
- b) Koska se luo perustan tiedolla johtamiselle ja strategiselle suunnittelulle.

- c) Koska sen avulla voidaan täyttää asiakastietokanta mahdollisimman nopeasti.
- d) Koska se vähentää tarvetta myyntiin ja asiakaskohtamisiin.

*Oikea vastaus: b) Datan kerääminen on liiketoiminnalle tärkeää, koska se tarjoaa perustan päätöksenteolle, tiedolla johtamiselle ja strategiselle suunnittelulle.*

**Organisaatiot voivat vahvistaa asiakkaiden luottamusta noudattamalla:**

- a) tietoturvaperiaatteita ja käyttämällä dataa vastuullisesti
- b) markkinointistrategioita ja hyödyntämällä dataa kaupallisesti
- c) IT-protokollia ja jakamalla dataa yhteistyökumppaneiden kanssa

*Oikea vastaus: a) Asiakkaiden luottamus perustuu siihen, että organisaatiot käsittelevät heidän tietojaan eettisesti ja turvallisesti. GDPR edellyttää, että organisaatiot noudattavat selkeitä tietosuojaperiaatteita ja käyttävät dataa vastuullisesti liiketoiminnassaan.*

**Mikä seuraavista EI kuulu GDPR:n keskeisiin periaatteisiin?**

- a) Datan keräämisen tarpeellisuus
- b) Asiakkaiden tietojen anonymisointi
- c) Datan myynti kolmansille osapuolille ilman suostumusta
- d) Läpinäkyvyys tiedon käytössä

*Oikea vastaus: c) GDPR:n keskeisiin periaatteisiin kuuluvat datan keräämisen tarpeellisuus, asiakkaiden tietojen anonymisointi ja läpinäkyvyys tiedon käytössä. Datan myynti kolmansille osapuolille ilman suostumusta rikkoo GDPR:n määräyksiä, sillä henkilötietojen luovuttaminen edellyttää aina selkeää ja yksiselitteistä suostumusta henkilöltä, jonka tietoja käsitellään. Tämä periaate on olennainen yksilöiden oikeuksien suojaamiseksi ja luottamuksen säilyttämiseksi.*

## Tehtävä 6

Essee

Miten GDPR voi auttaa organisaatiota vahvistamaan asiakkaiden luottamusta? Kirjoita vastaus, jossa pohdit suostumuksen hallinnan, läpinäkyvyyden ja tietoturvan merkitystä asiakassuhteissa.

# Datan säilytys ja hallinta

Kun data on kerätty, sen asianmukainen säilytys on välttämätöntä sen varmistamiseksi, että data on käyttökelpoista, turvallista ja saatavilla. Datan säilytyksellä tarkoitetaan tietojen tallentamista järjestelmällisesti niin, että ne voidaan myöhemmin hakea analysointia ja hyödyntämistä varten. Ilman luotettavaa säilytysjärjestelmää data voi kadota tai muuttua virheelliseksi tai käyttökeltottomaksi. Säilytys toimii pohjana kaikille datan käsittelyyn liittyville prosesseille – analysoinnista raportointiin ja päätöksentekoon.

## Opi lisää

### Säilytyksen periaatteet

Jotta data olisi käyttökelpoista ja turvallista, säilytyksessä tulee huomioida:

- **Tietoturva:** Henkilötietojen suojaamiseen ja tietovuotojen estämiseen liittyvät prosessit. GDPR-asetus korostaa, että henkilötiedot on säilytettävä suojatussa ympäristössä.
- **Eheys:** Datan tulee säilyä muuttumattomana tallennuksen aikana.
- **Saatavuus:** Tallennettu data on oltava helposti ja nopeasti käytettävissä.
- **Skaalautuvuus:** Säilytysjärjestelmien on kyettävä mukautumaan organisaation kasvaviin tarpeisiin.

Hyvin toteutettu säilytys mahdollistaa:

- **tietoturvan takaamisen:** organisaation ja asiakkaiden tiedot suojataan luvattomalta käytöltä
- **nopean saatavuuden:** data on käytettävissä reaaliaikaisiin tarpeisiin
- **laadun ylläpidon:** säilytysjärjestelmät pitävät tiedot eheinä ja luotettavina.

### Säilytyksen teknologiat ja ratkaisut

Datan säilytysteknologiat ovat kehittyneet merkittävästi viime vuosikymmeninä, ja nykyään organisaatioilla on käytössään monia vaihtoehtoja tallentaa ja hallita tietoa tehokkaasti. Säilytysratkaisujen valinta riippuu pitkälti organisaation tarpeista, budjetista ja tietoturva-vaatimuksista. Käydään seuraavaksi läpi paikallisia järjestelmiä, pilvipohjaisia ratkaisuja, hajautettuja tietokantajärjestelmiä ja hybridiratkaisuja datan säilytykseen.

## **Paikalliset järjestelmät**

Paikalliset datakeskukset ja palvelimet ovat perinteinen tapa tallentaa dataa. Ne tarjoavat organisaatiolle täydellisen hallinnan tiedoista, mutta vaativat suuria alkuinvestointeja ja jatkuvaa ylläpitoa. Paikalliset järjestelmät ovat erityisen sopivia erittäin arkaluonteista tietoa käsitteleville organisaatioille esimerkiksi rahoitus- tai terveydenhuollon aloilla.

### **Edut:**

- täysi hallinta tiedoista ja infrastruktuurista
- ei riippuvuutta ulkoisista palveluntarjoajista.

### **Haitat:**

- korkeat kustannukset infrastruktuurin ylläpidossa ja laajentamisessa
- vähemmän joustavuutta verrattuna pilvipohjaisiin ratkaisuihin.

### **Esimerkki:**

Sairaalat ja terveydenhuollon organisaatiot käyttävät usein paikallisia järjestelmiä potilastietojen säilyttämiseen. Koska potilastiedot ovat äärimmäisen arkaluonteisia, niiden käsittelyyn liittyy tiukkoja tietoturva- ja tietosuojavaatimuksia, kuten edellisessä kappaleessa esittelemämme GDPR. Paikalliset järjestelmät mahdollistavat täyden hallinnan tietojen käyttöön ja säilytykseen liittyvissä prosesseissa. Esimerkiksi leikkausten aikataulut, laboratoriotulokset ja potilashistoriatiedot tallennetaan turvallisiin datakeskuksiin, joihin vain valtuutetuilla käyttäjillä on pääsy.

Tällainen ratkaisu minimoi ulkoisten tietovuotojen riskit ja varmistaa järjestelmän käytettävyyden myös silloin, kun internet-yhteyksissä ilmenee häiriöitä. Haittapuolena ovat korkeat ylläpitokustannukset ja infrastruktuurin päivittämisen vaatima IT-osaaminen.

## **Pilvipohjaiset ratkaisut**

Pilvipalvelut, kuten Amazon Web Services (AWS), Google Cloud ja Microsoft Azure, ovat mullistaneet datan säilytyksen tarjoamalla skaalautuvia, kustannustehokkaita ja helposti saatavilla olevia ratkaisuja. Pilvipalvelut mahdollistavat datan tallentamisen ja hakemisen mistä tahansa, mikä tukee liiketoiminnan jatkuvuutta ja etätyöskentelyä. Skaalautuvuudella tarkoitetaan myös sitä, että pilveä voi laajentaa liiketoiminnan ja tarpeiden kasvaessa tai päinvastoin pienentää tarvittaessa mahdollistaen kustannustehokkaamman datan keruun.

**Edut:**

- Skaalautuvuus: kapasiteetti kasvaa tai vähenee tarpeen mukaan ilman suuria alkuinvestointeja.
- Joustavuus: tietoihin pääsee käsiksi mistä tahansa, mikä tekee järjestelmistä ihanteellisia etätyöhön ja hajautettuihin tiimeihin.
- Tietoturva: palveluntarjoajat tarjoavat edistyneitä tietoturvaratkaisuja, kuten salauksen ja käyttöoikeuksien hallinnan.

**Haitat:**

- Riippuvuus ulkoisista palveluntarjoajista, mikä voi aiheuttaa haasteita esimerkiksi palvelukatkosten aikana.
- Potentiaalinen tietoturvariski: vaikka pilvipalvelut tarjoavat vahvaa tietoturvaa, ulkoiset uhat, kuten tietomurrot, ovat mahdollisia.
- Oikeudelliset ja sijaintiin liittyvät kysymykset: dataa voidaan säilyttää palvelimilla eri maissa, mikä voi aiheuttaa haasteita tietosuojalakien, kuten GDPR:n, noudattamisessa.
- Vaatii internet-yhteyden: pilvipalvelut eivät toimi ilman luotettavaa internet-yhteyttä, mikä voi olla ongelma alueilla, joilla yhteys on hidaskä tai epäluotettava.

**Esimerkki:**

Verkkokaupat, jotka käsittelevät suuria määriä asiakasdataa, hyödyntävät usein pilvipalveluja varmistaakseen datan nopean saatavuuden, turvallisuuden sekä kustannustehokkuuden suuren kysynnän aikana.

**Hajautetut tietokantajärjestelmät**

Hajautetut järjestelmät, kuten Apache Cassandra ja MongoDB, mahdollistavat datan tallentamisen ja hallinnan useissa eri sijainneissa samanaikaisesti. Näin järjestelmät pysyvät toiminnassa, vaikka yksi palvelin vikaantuisi. Tämä tekee niistä erityisen hyödyllisiä kriittisille sovelluksille, joissa jatkuvuus on tärkeää.

**Edut:**

- Korkea käytettävyys: data on saatavilla, vaikka yksi tai useampi palvelin vikaantuisi.
- Nopea tiedonhaku: suurten datamäärien hallinta tehokkaasti reaaliajassa.
- Kestävyys: hajautettu rakenne estää yksittäistä vikaa kaatamasta koko järjestelmää.

**Haitat:**

- Monimutkaisuus: hajautetut järjestelmät voivat olla teknisesti haastavia suunnitella ja ylläpitää.
- Korkeammat kustannukset verrattuna yksinkertaisempiin järjestelmiin, erityisesti alkuvaiheessa.
- Datan eheysongelmat: ristiriitaisuudet eri solmujen välillä voivat johtaa tiedon vääristymiseen.

**Esimerkki:**

Reaaliaikaiset analyysijärjestelmät, kuten mobiilipelit ja suoratoistopalvelut, käyttävät hajautettuja tietokantoja, jotta ne voivat käsitellä käyttäjätietoja nopeasti ja luotettavasti.

**Hybridiratkaisut**

Hybridiratkaisut yhdistävät paikalliset järjestelmät ja pilvipalvelut. Tämä mahdollistaa arkaluonteisten tietojen säilyttämisen paikallisesti ja muun datan joustavan hallinnan pilvessä. Hybridiratkaisut ovat erityisen suosittuja suurissa organisaatioissa, joissa on monimutkaiset säilytysvaatimukset.

**Edut:**

- Joustavuus: mahdollisuus hyödyntää paikallisten järjestelmien ja pilvipalveluiden parhaat puolet.
- Arkaluonteisten tietojen säilyttäminen paikallisesti, mutta yleisten analyysien ja raporttien hallinta pilvessä.

**Haitat:**

- Monimutkaisuus: hybridiratkaisujen suunnittelu ja hallinta vaativat huomattavia resursseja ja asiantuntemusta.
- Korkeammat kustannukset, koska kahden eri järjestelmän ylläpidosta aiheutuu yhteensä suuremmat kulut.

**Esimerkki:**

Pankki voi säilyttää asiakkaiden henkilötietoja paikallisesti tiukan tietoturvan takaamiseksi, mutta käyttää pilvipalveluja yleisten analyysien ja raporttien tuottamiseen.

Datan säilytys on kriittinen osa tiedolla johtamista, sillä se luo perustan datan myöhemmälle analysoinnille ja hyödyntämiselle. Säilytysratkaisuissa on tärkeää huomioida tietoturva, eheys ja saatavuus, jotta data on käyttökelpoista kaikissa organisaation prosesseissa. Oikein toteutettu säilytysratkaisu ei ainoastaan tue päivittäistä toimintaa, vaan myös parantaa organisaation kykyä mukautua tulevaisuuden tarpeisiin.

## Tehtävä 7

Valitse oikea vastaus.

### **Mikä on datan säilytyksen keskeinen tavoite?**

- a) Datan muuttaminen käyttökelvottomaksi ajan myötä.
- b) Varmistaa, että data on käyttökelpoista, turvallista ja helposti saatavilla.
- c) Estää datan jakaminen eri järjestelmien välillä.
- d) Säilyttää data ainoastaan paikallisilla palvelimilla.

*Oikea vastaus: b) Datan säilytys on tärkeää sen varmistamiseksi, että tieto säilyy käyttökelpoisena ja turvallisena analysointia ja päätöksentekoa varten.*

### **Mikä on paikallisten datakeskusten keskeinen etu?**

- a) Ne mahdollistavat tiedon helpon jakamisen julkisesti.
- b) Ne ovat aina kustannustehokkaampi vaihtoehto kuin pilvipalvelut.
- c) Ne vähentävät tarvetta IT-osaamiselle ja infrastruktuurin ylläpidolle.
- d) Ne tarjoavat organisaatiolle täyden hallinnan datasta ja infrastruktuurista.

*Oikea vastaus: d) Paikalliset datakeskukset antavat organisaatioille täyden hallinnan tietoturvaan ja datan säilytykseen, mikä on erityisen tärkeää arkaluonteisessa datassa.*

### **Mikä on pilvipalveluiden etu verrattuna paikallisiin järjestelmiin?**

- a) Ne tarjoavat täydellisen hallinnan tiedoista ja infrastruktuurista.
- b) Ne ovat täysin ilmaisia käyttää.
- c) Ne tarjoavat skaalautuvuutta ja joustavuutta, kuten mahdollisuuden käyttää dataa eri sijainneista ja mukauttaa resurssit tarpeen mukaan.
- d) Ne eivät vaadi tietoturvatouimia.

*Oikea vastaus: c) Pilvipalvelut tarjoavat skaalautuvuutta ja joustavuutta, mikä mahdollistaa datan käytön eri sijainneista ja resurssien mukauttamisen tarpeiden mukaan.*

### **Mikä seuraavista on hajautetun tietokantajärjestelmän etu?**

- a) Se mahdollistaa datan nopean haun ja korkean käytettävyyden.
- b) Se vähentää tarvetta varmuuskopioinnille.
- c) Se vaatii vähemmän teknistä asiantuntemusta kuin muut järjestelmät.
- d) Se toimii vain paikallisilla palvelimilla ilman internet-yhteyttä.

*Oikea vastaus: a) Hajautetut tietokantajärjestelmät takaavat korkean käytettävyyden, sillä ne pystyvät toimimaan useilla palvelimilla samanaikaisesti.*

### **Miksi organisaatio voisi valita hybridiratkaisun datan säilytykseen?**

- a) Se yhdistää paikallisten järjestelmien ja pilvipalveluiden parhaat puolet.
- b) Se poistaa tarpeen suojata arkaluonteista tietoa.
- c) Se on aina halvin vaihtoehto datan säilyttämiseen.
- d) Se mahdollistaa datan säilyttämisen ilman internet-yhteyttä.

*Oikea vastaus: a) Hybridiratkaisut mahdollistavat arkaluonteisen tiedon säilyttämisen paikallisesti samalla, kun vähemmän kriittistä dataa hallinnoidaan pilvipalveluissa.*

## **Datan mallintaminen**

Datan mallintaminen tarkoittaa tietojen järjestämistä loogiseen rakenteeseen, jotta niitä voidaan analysoida ja hyödyntää tehokkaasti. Mallintamisen avulla raakadatasta saadaan selkeää, ymmärrettävää ja liiketoiminnan tarpeisiin soveltuvaa tietoa.

Yksi yleisimmistä tavoista mallintaa dataa on jakaa se faktatauluihin ja dimensiotauluihin. Faktataulu sisältää tapahtumatason tietoa, kuten ostotapahtumat, asiakaskäynnit tai verkkosivujen klikkaukset, kun taas dimensiotaulut tarjoavat kontekstin näille tapahtumille.

Mallintamisen tavoitteena on tehdä datasta:

- ymmärrettävää: tietojen suhteet ovat selkeitä ja helposti tulkittavia
- joustavaa: eri näkökulmien analysointi onnistuu ilman monimutkaisia laskutoimituksia
- nopeaa käsitellä: tietokannat voivat optimoida kyselyjä tehokkaammin.

Kun data on mallinnettu hyvin, yritykset voivat tehdä parempia päätöksiä ja hyödyntää analytiikkaa esimerkiksi asiakaskäyttäytymisen ymmärtämiseen, myynnin optimointiin ja resurssien tehokkaaseen kohdentamiseen.

Käsitellään seuraavaksi datan mallintamista käytännön tasolla. Kun asiakas ostaa kaupasta tuotteen, siitä syntyy tapahtuma eli rivi tietokantaan jota kutsutaan faktaksi. Faktarivi sisältää esimerkiksi ostetun määrän ja hinnan, mutta ei välttämättä suoraan kerro tuotteen nimeä. Sen sijaan siinä on tuotenumero, joka toimii avaimena lisätietoihin. Tarkemmat tiedot tuotteesta, kuten nimi, kategoria ja brändi, löytyvät erillisestä dimensiotaulusta, joka on tässä esimerkissä kaupan tuotetaulu. Tämä mahdollistaa myyntidatan analysoinnin eri näkökulmista, kuten tuoteryhmittäin, brändeittäin tai hintaluokittain.

Tätä periaatetta voidaan soveltaa myös muihin tietoihin. Esimerkiksi asiakasnumero voi olla faktarivillä, mutta asiakkaan nimi ja asiakassegmentti löytyvät erillisestä asiakastaulusta. Samoin myymälän nimi ja sijainti eivät ole suoraan faktarivillä, vaan ne haetaan myymälätiedoista.

Tapahtumatasoinen eli **transaktiotason data** tallentaa yksittäiset ostot, maksut ja muut tapahtumat tarkimmalla tasolla. Sen avulla voidaan tehdä syvällistä analyysiä, kuten:

- myynnin seuraaminen tuoteryhmätasolla: faktoja ja dimensioita yhdistämällä voidaan tarkastella, mitkä tuotteet myyvät eniten.
- poikkeamien tunnistaminen: jos myyntiraportissa näkyy epätavallinen piikki, voidaan porautua yksittäisiin ostoksiin ja selvittää syy.
- trendien ymmärtäminen: yksittäisten ostojen perusteella voidaan havaita asiakkaiden käyttäytymisen muutoksia.

Kun data on mallinnettu faktojen ja dimensioiden avulla, siitä tulee helposti käytettävää ja analysoitavaa eri näkökulmista. Tämä rakenne tekee tiedosta selkeämpää ja tukee parempaa päätöksentekoa.

## Esimerkki

### Käytännön esimerkki: hintakampanjan vaikutuksen analysointi

Supermarket-ketju haluaa arvioida viikonloppukampanjansa vaikutusta, jossa tiettyjä tuotteita myydään alennettuun hintaan. Ilman tapahtumatasoista dataa voidaan nähdä vain, että jogurttien myynti kasvoi 20 %, mutta ei voida varmuudella sanoa, johtuiko kasvu kampanjasta vai muista tekijöistä.

Tapahtumadata mahdollistaa ostoskorianalyysin, jossa tarkastellaan, mitä muita tuotteita asiakkaat ostivat alennustuotteen lisäksi. Jos asiakkaat ostivat pelkästään tarjoustuotteita, kampanja ei välttämättä kasvattanut kokonaistuottoja. Lisäksi voidaan selvittää, perustuiko myynnin kasvu uusien asiakkaiden hankintaan vai kanta-asiakkaiden ostotiheyden muutoksiin.

Taloudellisen kannattavuuden arvioimiseksi voidaan tutkia, vaihtoivatko asiakkaat

normaalihintaiset premium-tuotteet edullisempiin alennustuotteisiin. Jos näin tapahtui laajasti, myynnin kasvu saattoi tulla katemarginaalin kustannuksella.

Ilman tapahtumatason dataa johtopäätökset perustuivat vain summatason myyntilukuihin, mikä voi johtaa harhaan. Summatasolla voidaan nähdä kokonaismyynnin kasvu, mutta ilman tarkempaa tietoa yksittäisistä ostoksista ei tiedetä, mitä asiakkaat todella tekivät ja mikä kampanjan vaikutus oli heidän ostokäyttäytymiseensä.

### **Mitä voi tapahtua, jos tapahtumatason data puuttuu?**

Esimerkiksi:

- Kampanjan jälkeen havaitaan, että jogurttien myynti nousi 20 %, joten päätetään toistaa kampanja kuukausittain.
- Ilman tapahtumadataa ei kuitenkaan huomata, että suurin osa ostajista oli kanta-asiakkaita, jotka normaalisti ostavat kalliimpia tuotteita, mutta siirtyivät nyt alennustuotteisiin.
- Kun kampanja toistetaan, premium-tuotteiden myynti laskee pysyvästi, ja vaikka volyymi kasvaa, kokonaistuotto pienenee.
- Myös muiden tuotteiden myynti laskee, koska asiakkaat ostivat vain tarjoustuotteita eivätkä täydentäneet ostoskorja muilla elintarvikkeilla.

Tämä johtaa siihen, että kampanja näyttää paperilla onnistuneelta, mutta pitkällä aikavälillä se heikentää myyntikatteita eikä asiakaskäyttäytymistä ymmärretä oikein.

Tarkempi analyysi auttaa kehittämään tehokkaampia kampanjoita, jotka kannustavat asiakkaita ostamaan myös muita tuotteita ja parantavat myynnin kannattavuutta.

## Tehtävä 8

Valitse oikea vastaus.

**Datan mallintaminen tarkoittaa ainoastaan sen tallentamista eri tietokantoihin ilman, että tietojen välisiä suhteita määritellään.**

- a) Oikein
- b) Väärin

*Oikea vastaus: Väärin. Datan mallintaminen sisältää tietojen rakenteen ja suhteiden määrittämisen, mikä mahdollistaa tehokkaan analyysin ja nopean tiedonhaun. Esimerkiksi faktojen ja dimensioiden käyttäminen auttaa luomaan selkeän tietomallin.*

### **Mikä on datan mallintamisen keskeinen tavoite?**

- a) Vähentää datan määrää poistamalla tarpeettomat tiedot.
- b) Järjestää data loogiseen rakenteeseen, jotta sitä voidaan analysoida ja hyödyntää tehokkaasti.
- c) Estää datan yhdistäminen eri lähteistä.
- d) Tallentaa kaikki tiedot yhteen tauluun ilman rakenteita.

*Oikea vastaus: b) Datan mallintaminen tarkoittaa tietojen järjestämistä loogiseen rakenteeseen, jotta niitä voidaan analysoida ja hyödyntää tehokkaasti.*

### **Mikä erottaa faktataulut ja dimensiotaulut toisistaan?**

- a) Faktataulut sisältävät tapahtumatason tietoa, kun taas dimensiotaulut tarjoavat kontekstin näille tapahtumille.
- b) Faktataulut sisältävät vain tuotteiden nimet ja hinnat, kun taas dimensiotaulut tallentavat tapahtumatason tiedot.
- c) Dimensiotaulut ovat aina suurempia kuin faktataulut.
- d) Faktataulut ja dimensiotaulut ovat sama asia ja niitä käytetään samalla tavalla.

*Oikea vastaus: a) Faktataulut sisältävät tapahtumatason tietoa, kuten ostotapahtumat, kun taas dimensiotaulut tarjoavat näille tapahtumille kontekstin, kuten tuotteen tai asiakkaan tiedot.*

## **Datan analysointi**

Datan analysointi on prosessi, jossa kerätystä tiedosta jalostetaan hyödyllistä tietoa päätöksenteon tueksi. Tämä on kriittinen osa tiedolla johtamista, sillä ilman analysointia data jää raaka-aineeksi, jota ei voida hyödyntää tehokkaasti. Analysoinnin tavoitteena on tuottaa konkreettisia oivalluksia, jotka auttavat organisaatioita tekemään tietoon perustuvia päätöksiä, parantamaan suorituskykyään ja löytämään uusia liiketoimintamahdollisuuksia. Analyysin tuottama tieto voi paljastaa trendejä, parantaa tehokkuutta ja tukea strategista suunnittelua.

### **Käsite**

#### **Analysointi ja analytiikka**

Analyysi ja analytiikka ovat läheisesti toisiinsa liittyviä, mutta ne viittaavat eri asioihin.

### **Analyysi:**

- kohdistuu yksittäiseen ongelmaan tai kysymykseen
- voi olla kertaluontoinen, kuten myyntidatan tarkastelu yhden kuukauden ajalta
- käyttää usein perinteisiä tilastollisia menetelmiä tai yksinkertaisia työkaluja
- vastaa kysymyksiin "Mitä tapahtui?" ja "Miksi näin tapahtui?"
- Esimerkki: Yritys analysoi tietyn markkinointikampanjan tulokset ja havaitsee, että sen ROI, eli sijoitetun pääoman tuotto (englanniksi *Return on Investment*) jäi odotettua alhaisemmaksi.

### **Analytiikka:**

- on järjestelmällinen ja jatkuva prosessi, jossa dataa hyödynnetään laajasti
- käyttää kehittyneitä teknologioita, kuten tekoälyä ja automaatiojärjestelmiä
- pyrkii ennakoimaan tulevaa ja antamaan suosituksia päätöksenteon tueksi
- vastaa kysymyksiin "Mitä tapahtuu seuraavaksi?" ja "Mitä meidän pitäisi tehdä?"
- Esimerkki: Verkkokauppa käyttää analytiikkaa ennustamaan asiakaskäyttäytymistä ja mukauttaa varastonhallintansa ennusteiden perusteella.

## **Analytiikan neljä tasoa**

Datan analysointi voidaan jakaa neljään pääasialliseen lähestymistapaan, jotka vastaavat eri liiketoimintatarpeisiin:

### **1. Kuvaileva analytiikka eli raportointi**

Kuvaileva analytiikka vastaa kysymykseen *Mitä tapahtui?* Se perustuu historiallisen datan tarkasteluun ja auttaa ymmärtämään menneitä tapahtumia. Esimerkiksi myyntitilastojen analysointi paljastaa trendejä ja sesonkivaihteluita.

### **2. Diagnosoiva analytiikka ja interaktiivisuus**

Diagnosoiva analytiikka syventää kuvailevaa analytiikkaa vastaamalla kysymykseen *Miksi näin tapahtui?* Tämä lähestymistapa keskittyy syy-seuraussuhteiden löytämiseen, esimerkiksi selvittämällä, miksi tietyn tuotteen myynti on laskenut.

### **3. Ennakoiva analytiikka**

Ennakoiva analytiikka käyttää historiallista dataa ja malleja vastatakseen kysymykseen *Mitä todennäköisesti tapahtuu?* Esimerkiksi ennustemallit voivat arvioida, miten markkinat muuttuvat seuraavan vuosineljänneksen aikana.

## 4. Ohjaava analytiikka

Edistynein analytiikan muoto, ohjaava analytiikka, vastaa kysymykseen *Mitä pitäisi tehdä?* Tämä taso hyödyntää tekoälyä ja optimointimalleja tarjotakseen konkreettisia suosituksia päätöksenteon tueksi.

### Perinteinen datan analysointi: tilastolliset menetelmät ja käsin tehtävä analyysi

Tilastolliset tunnusluvut, kuten keskiarvot, mediaanit, varianssi ja korrelaatiot, tarjoavat nopean tavan tunnistaa peruspiirteitä datasta. Tilastollisilla menetelmillä, kuten liukuvalla keskiarvolla tai regressioanalyysillä voidaan ennustaa esimerkiksi asiakaskäyttötymisen yleisiä trendejä, tuotekohtaisia myyntipiikkejä tai markkinamuutoksia.

Esimerkiksi vähittäiskaupassa yritys voi analysoida asiakasmääriä ja vertailla niitä eri viikkojen välillä ymmärtääkseen sesonkivaihteluita. Vastaavasti finanssisektorilla käytetään perinteisiä aikasarja-analyysijä, joiden avulla voidaan arvioida esimerkiksi osakkeiden hintakehitystä tai korkotason vaikutuksia markkinoihin.

Vaikka perinteiset analyysimenetelmät ovat edelleen tärkeitä, niiden suurin rajoite on skaalautuvuus. Kun datamäärät kasvavat suuriksi, manuaalinen käsittely ja tilastolliset mallit eivät aina riitä löytämään monimutkaisia yhteyksiä tai tekemään tarkkoja ennusteita.

### Analysointiprosessin neljä vaihetta

#### 1. Määrittele tavoite

Analyyysin lähtökohtana on ymmärtää, mitä kysymyksiä halutaan ratkaista. Esimerkiksi:

- Mitkä tuotteet ovat myyneet parhaiten viime kuussa?
- Mikä on asiakaskunnan yleisin valitusaihe? Selkeä tavoite ohjaa analyyysin fokusta ja auttaa valitsemaan sopivat työkalut ja menetelmät.

#### 2. Valmistele data

Ennen analysointia on varmistettava, että käytössä oleva data on luotettavaa ja yhtenäistä. Valmisteluvaiheessa:

- tunnistetaan ja korjataan virheelliset tai puutteelliset tiedot
- yhtenäistetään dataa, esimerkiksi muuntamalla valuutat tai aikamuodot yhdenmukaisiksi.

### 3. Analysoi data

Valitse analyysimenetelmä tavoitteiden ja datan perusteella. Esimerkkejä menetelmistä:

- laske keskiarvot tai mediaanit saadaksesi yleiskuvan.
- suodata dataa segmenttikohtaisesti, kuten asiakasryhmittäin, tunnistaaksesi tarkempia trendejä.

### 4. Tulkitse tulokset

Pohdi, mitä löydökset kertovat ja miten niitä voidaan hyödyntää. Esimerkiksi:

- Jos tietty tuote on myynyt erityisen hyvin, pitäisikö markkinointia kohdentaa sen ympärille?

### Perusmenetelmät analysointiin

Käytännön analysoinnissa voidaan käyttää seuraavia menetelmiä:

- **Suodattaminen ja lajittelu**  
Esimerkiksi Excelissä tai Google Sheetsissä suodattamalla tietoja, kuten myyntimääriä tai asiakassegmenttejä, voidaan tunnistaa trendejä.
- **Ryhmittely ja tiivistäminen**  
Ryhmittelemällä dataa esimerkiksi alueittain, iän mukaan tai tuoteryhmittäin voidaan havaita toistuvia ilmiöitä.
- **Tunnusluvut**  
Yksinkertaiset tilastolliset mittarit, kuten keskiarvo, mediaani tai varianssi, tarjoavat nopean tavan ymmärtää datan perusominaisuuksia.

### Työkalut analyysin tueksi

Datan analysointiin on saatavilla laaja valikoima työkaluja, jotka sopivat eri taitotasolle. Tässä muutamia työkaluja joita tyypillisesti käytetään analysointityössä:

- **Excel ja Google Sheets:** sopivat aloittelijoille. Näiden avulla voidaan laskea tunnuslukuja, tehdä peruskaavioita ja käyttää suodattimia.
- **Power BI, Qlik ja Tableau:** tarjoavat visuaalisempia ja monipuolisempia tapoja analysoida suuria tietomääriä ja luoda raportteja.

- **Koodauskielet Python ja R:** sopivat edistyneempiin tarpeisiin. Näillä ohjelmointikielillä voidaan toteuttaa monimutkaisempia analyyskejä, kuten koneoppimismalleja tai tilastollisia simulaatioita.

## Tehtävä 9

Valitse oikea vaihtoehto.

### Mitä eroa on analyysillä ja analytiikalla?

- Analyyysi on järjestelmällinen ja jatkuva prosessi, kun taas analytiikka on kertaluontoinen tapahtuma.
- Analyyysi keskittyy menneisiin tapahtumiin ja niiden syihin, kun taas analytiikka pyrkii ennakoimaan tulevaa ja tarjoamaan suosituksia.
- Analyyysi käyttää kehittyneitä teknologioita, kuten tekoälyä, kun taas analytiikka perustuu pääasiassa perinteisiin tilastollisiin menetelmiin.
- Analyyysi ja analytiikka tarkoittavat samaa asiaa, mutta niitä käytetään eri konteksteissa.

*Oikea vastaus: b) Analyyysi vastaa kysymyksiin "Mitä tapahtui?" ja "Miksi näin tapahtui?", ja se voi olla kertaluontoinen tarkastelu. Analytiikka puolestaan on jatkuva prosessi, joka hyödyntää kehittyneitä teknologioita, kuten tekoälyä, ja vastaa kysymyksiin "Mitä tapahtuu seuraavaksi?" ja "Mitä meidän pitäisi tehdä?".*

### Mikä seuraavista on esimerkki ennakoivasta analytiikasta?

- Yritys analysoi viime vuoden myyntilukuja ja havaitsee, että joulukuu oli paras myyntikuukausi.
- Yritys tutkii, miksi tietyn tuotteen myynti laski merkittävästi viime kuussa.
- Verkkokauppa käyttää asiakasdataa ennustukseen, mitkä tuotteet tulevat olemaan suosittuja seuraavan sesongin aikana.
- Yritys laskee tuoteryhmän keskiarvoisen myyntimäärän päivittäin Excelissä.

*Oikea vastaus: c) Ennakoiva analytiikka hyödyntää historiallista dataa ja malleja ennustukseen tulevaisuuden tapahtumia. Esimerkissä verkkokauppa pyrkii arvioimaan tulevia suosikkituotteita, mikä on tyypillinen ennakoivan analytiikan käyttötapa. Muut vaihtoehdot liittyvät enemmän kuvailevaan (a, d) tai diagnosoivaan (b) analytiikkaan.*

# Tekoälyyn perustuva analytiikka

Siinä missä perinteinen analytiikka keskittyy kuvailemaan tapahtuneita asioita, edistynyt, eli tekoälyyn perustuva analytiikka (englanniksi *advanced analytics*) ennustaa tulevaa ja antaa konkreettisia suosituksia päätöksenteon tueksi. Vaikka iso osa datan käsittelystä tehdään edelleen hartiavoimin, tekoäly on viime vuosina mullistanut analytiikan mahdollisuuksia merkittävästi. Edistynyt analytiikka liittyy aiemmin esiteltyyn analytiikan taso -jaotteluun siten, että se kattaa ennakoivan ja ohjaavan analytiikan, joissa hyödynnetään edistyneitä menetelmiä, kuten tekoälyä, koneoppimista ja tilastollisia malleja.

## Käsite

### Edistyneen analytiikan keskeiset käsitteet

#### Koneoppiminen:

- **Valvottu oppiminen:** Malli oppii ennustamaan tuloksia perustuen merkittyihin tietoihin. Esimerkiksi asiakaskäyttäytymisen ennustaminen.
- **Valvomaton oppiminen:** Malli löytää piilossa olevia rakenteita datasta ilman ennalta määriteltäviä tuloksia. Esimerkiksi asiakassegmentointi.
- **Puoliksi valvottu oppiminen:** Yhdistelmä valvottua ja valvomattonta oppimista.

#### Syväoppiminen:

- Koneoppimisen osa-alue, joka käyttää tehokkaita neuroverkkoja. Neuroverkot ovat laskennallisia malleja, jotka jäljittelevät ihmisaivojen rakennetta ja toimintaa. Ne parantavat ennusteitaan kuljettamalla dataa useiden kerrosten eli neuronien läpi. Jokaisessa kerroksessa neuroverkko oppii ja parantaa tietoa, jonka avulla se pystyy tunnistamaan yhä tarkempia ja monimutkaisempia datakuvioita.
- Soveltuu erityisesti monimutkaisten ongelmien ratkaisemiseen suuria datamassoja hyödyntäen.
- Käytetään muun muassa kuvantunnistuksessa, luonnollisen kielen käsittelyssä ja lääketieteessä.

#### Luonnollisen kielen käsittely (NLP):

- Tietokoneiden kyky ymmärtää, tulkita ja tuottaa ihmiskieltä.
- Käytetään muun muassa asiakaspalveluchatboteissa, tekstianalyysissä ja sosiaalisen median seurannassa.

#### Tietojen louhinta:

- Löytää piilossa olevia kuvioita ja suhteita suurista datamääristä.

- Käytetään mm. markkinoinnissa, finanssialalla ja tieteellisessä tutkimuksessa.

Tekoälyn käyttö analytiikassa mahdollistaa monimutkaisempien mallien hyödyntämisen ja suurten datamassojen tehokkaan käsittelyn. Tekoäly ei ainoastaan nopeuta analyysia, vaan se voi myös havaita piileviä yhteyksiä ja ennakoida tulevia tapahtumia historiallisten tietojen perusteella.

Tekoälyn vaikutus datan analysointiin näkyy erityisesti kolmella osa-alueella: automaatioissa, ennustavassa analytiikassa ja päätöksenteon optimoinnissa. Organisaatiot käyttävät tekoälyn algoritmeja esimerkiksi suurten tietomassojen analysointiin, jolloin manuaalinen työ vähenee ja päätökset voidaan tehdä nopeasti ja johdonmukaisesti.

Tekoälyn suurin vahvuus on sen kyky mukautua ja oppia uusista tiedoista. Perinteisessä analytiikassa analyysi perustuu usein ennalta määriteltyihin kaavoihin ja mittareihin, kun taas koneoppimisen mallit ja tilastolliset menetelmät kehittyvät ajan myötä. Tämä tekee tekoälystä tehokkaan työkalun jatkuvasti muuttuvissa ympäristöissä. Esimerkiksi finanssialalla tekoäly voi havaita poikkeavuuksia, jotka perinteisillä menetelmillä saattaisivat jäädä huomaamatta, ja tunnistaa epäilyttäviä maksutapahtumia reaaliajassa.

Neuroverkkoihin perustuvat mallit kykenevät puolestaan käsittelemään suuria ja monimutkaisia tietojoukkoja, kuten asiakaspalautteita verkkokaupassa. Tekoäly pystyy analysoimaan käyttäjien klikkauksia, hakuhistoriaa ja ostokäyttäytymistä ja suosittelemaan tuotteita, jotka todennäköisesti kiinnostavat asiakasta.

Terveystieteiden tutkimuksessa tekoälypohjaiset analyysityökalut voivat havaita sairauksien varhaisia merkkejä potilastiedoista ja kuvantamistietojen avulla. Koneoppimisalgoritmit voivat tunnistaa esimerkiksi syöpäsolujen esiintymistä huomattavasti tarkemmin kuin perinteiset manuaaliset analyysit.

### **Kielimallien hyödyntäminen analytiikan tukena**

Tekoälypohjaiset kielimallit, kuten luonnollisen kielen käsittelyyn (NLP) perustuvat järjestelmät, tuovat analytiikkaan uuden ulottuvuuden, sillä ne kykenevät ymmärtämään, tulkitsemaan ja tuottamaan ihmiskieltä. Tämä mahdollistaa rakenteettoman datan tehokkaan käsittelyn sekä analyysin tulosten esittämisen ymmärrettävässä muodossa, mikä tekee analytiikasta helpommin lähestyttävää ja tukee datan laajempaa hyödyntämistä organisaation eri tasoilla.

Yksi keskeinen sovellusalue on tekstianalyysi, jossa kielimallit voivat käydä läpi asiakaspalautteita, tukipyyntöjä ja sosiaalisen median keskusteluja tunnistuen trendejä, tunteita ja

mahdollisia ongelmakohtia. Esimerkiksi asiakaspalvelussa tekoäly voi analysoida suuria määriä avointa palautetta ja nostaa esiin keskeisiä teemoja, jotka vaikuttavat asiakastyytyvyyteen.

Kielimalleja voidaan hyödyntää myös päätöksenteon tukena. Ne voivat käydä läpi laajoja tietovarantoja, kuten raportteja, tutkimuksia ja uutisvirtaa, tiivistäen olennaiset löydökset päätöksentekijöille. Tämä nopeuttaa tiedonhakua ja mahdollistaa parempien, faktoihin perustuvien päätösten tekemisen.

Automaattinen raportointi on toinen merkittävä sovellusalue. Kielimallit voivat tuottaa analytiikasta ymmärrettäviä yhteenvedoja ilman, että käyttäjän tarvitsee tulkita monimutkaisia tilastoja tai visualisointeja. Esimerkiksi myyntiraportti voidaan muuntaa selkeäksi tekstimuotoiseksi yhteenvedoksi, jossa nostetaan esiin keskeiset trendit ja poikkeamat.

Lisäksi kielimallit voivat tukea ennustavaa analytiikkaa yhdistämällä numeerista ja tekstimuotoista dataa. Esimerkiksi asiakaspalvelun yhteydenotoista voidaan tunnistaa, mitkä ongelmat ennakoivat asiakkaan poistumista (churn) ja millä toimenpiteillä asiakassuhde voitaisiin säilyttää.

## Tehtävä 10

Valitse oikea vaihtoehto.

### **Miten tekoäly parantaa datan analysointia verrattuna perinteisiin menetelmiin?**

- a) Tekoäly voi analysoida suuria datamääriä nopeasti ja tunnistaa piileviä yhteyksiä.
- b) Tekoäly korvaa täysin ihmisen tekemän päätöksenteon.
- c) Tekoäly tarvitsee aina käsin syötetyt säännöt toimiakseen.
- d) Tekoäly tekee vain yksinkertaista raportointia ilman ennustavia kykyjä.

*Oikea vastaus: a) Tekoäly ja koneoppiminen mahdollistavat suurten tietomassojen analysoinnin ja kykenevät löytämään monimutkaisia malleja ja yhteyksiä, joita perinteiset analyysimenetelmät eivät pysty tunnistamaan yhtä tehokkaasti.*

## Datan visualisointi

Datan analysoinnin tulokset voivat olla vaikeaselkoisia ilman asianmukaista esittämistapaa. Visualisoinnin avulla monimutkaisista tietomassoista voidaan luoda selkeitä ja helposti ymmärrettäviä esitysmuotoja, kuten kaavioita, diagrammeja ja interaktiivisia raportteja.

Visualisointi auttaa myös viestinnässä: johtoryhmät, tiimit ja muut sidosryhmät ymmärtävät paremmin datan pohjalta tehdyt päätökset ja oivallukset. Hyvin suunniteltu visualisointi ei ainoastaan tee tiedosta ymmärrettävää, vaan myös vakuuttavaa.

Käydään seuraavaksi läpi hyvän visualisoinnin periaatteet.

**1. Yksinkertaisuus ja selkeys:** visualisoinnin tulee olla mahdollisimman yksinkertainen ja helposti tulkittavissa. Liian monimutkaiset kaaviot tai liian monet elementit voivat tehdä datasta vaikeasti ymmärrettävää. Hyvä kaavio välittää vain olennaisen tiedon ilman häiriötekijöitä.

**2. Konteksti:** visualisoinnin on esitettävä dataa oikeassa kontekstissa. Tiedon esittäminen ilman taustatietoa voi hämmentää katsojaa. Esimerkiksi, jos esitetään myyntilukuja, on tärkeää kertoa myös, mihin ajanjaksoon tai alueeseen nämä luvut liittyvät.

**3. Värit ja muodot:** värit voivat vaikuttaa siihen, kuinka helposti kaavioita luetaan ja ymmärretään. Värien tulee olla selkeitä, erottuvia ja helposti tulkittavissa. Esimerkiksi tummat värit voivat edustaa suurempia arvoja, ja vaaleat värit pienempiä. Värien ei tule aiheuttaa sekaannusta tai olla liian häiritseviä.

**4. Interaktiivisuus:** erityisesti suurten ja monimutkaisempien datamäärien esittämisessä interaktiivisuus on usein erittäin hyödyllistä. Esimerkiksi interaktiiviset kaaviot ja dashboardit voivat antaa käyttäjälle mahdollisuuden tutkia tietoja tarkemmin, zoomata eri alueille ja tarkastella yksityiskohtaisia tietoja tarpeen mukaan. Tämä parantaa analyysin syvyyttä ja antaa käyttäjälle enemmän hallintaa tiedon tarkastelussa.

**5. Tarina ja oivallukset:** hyvä visualisointi ei ole pelkästään datan esittämistä, vaan sen avulla kerrotaan myös tarina. Visualisoinnin tulisi tarjota katsojalle mahdollisuus tehdä oivalluksia ja nähdä, mitä data kertoo. Tämä on erityisen tärkeää liiketoiminnan ja päätöksenteon kannalta, koska se auttaa ymmärtämään, miksi tiettyihin toimiin on ryhdyttävä.

**6. Tarkoituksenmukaisuus:** on tärkeää valita oikea visualisointityyppi, joka parhaiten palvelee tiedon esittämisen tarkoitusta. Esimerkiksi, jos halutaan vertailla eri kategorioiden osuus kokonaisuudesta, piirakkakaavio voi olla oikea valinta. Jos taas halutaan tutkia trendiä ajan kuluessa, viivakaavio on tehokas.

Erilaiset visualisointimenetelmät palvelevat erilaisia käyttötarkoituksia:

- **Viivakaaviot** sopivat trendien ja aikasarjojen esittämiseen, kuten myynnin kehittymisen seuraamiseen kuukausittain.
- **Pylväs- ja ympyrädiagrammit** auttavat vertailemaan eri muuttujia keskenään ja näyttämään prosentiosuuksia.

- **Lämpökartat (englanniksi *heatmap*)** visualisoivat tietyn ilmiön tiheyttä tai voimakkuutta, kuten verkkosivun eri osien suosiota käyttäjäseurannassa.
- **Interaktiiviset dashboardit** eli kokoavat näkymät yhdistävät useita eri visualisointeja yhdeksi kojelautaa muistuttavaksi kokonaisuudeksi. Ne tarjoavat käyttäjille mahdollisuuden suodattaa ja tarkastella tietoa eri näkökulmista.

Esimerkiksi johtoryhmä, joka tekee strategisia päätöksiä, tarvitsee datan esittämiseen **yhteenvedettyjä visualisointeja**, jotka tiivistävät keskeiset trendit ja havainnot. Analytikoille taas voi olla hyödyllistä nähdä **yksityiskohtaisempia visualisointeja**, joiden avulla voidaan syventyä datan taustalla oleviin tekijöihin.

## Opi lisää

### Tekoälyavusteiset työkalut visualisoinnin avuksi

Tekoäly voi myös auttaa visualisoinnin optimoinnissa. Nykyisin on olemassa tekoälyavusteisia työkaluja, jotka analysoivat dataa ja ehdottavat parhaita tapoja sen esittämiseen. Esimerkiksi Microsoft Power BI, Qlik ja Tableau käyttävät tekoälyä suosittelemaan käyttäjälle relevantteja kaaviotyyppejä ja havaintojen esittämistapoja.

## Keskeisimmät visualisoinnin virheet ja niiden välttäminen

Visualisointi tekee datasta helpommin ymmärrettävää, mutta väärin toteutettuna se voi myös johtaa harhaan ja aiheuttaa virheellisiä tulkintoja. Käydään seuraavaksi läpi yleisimmät virheet ja keinot niiden välttämiseen.

- **Mittasuhteiden vääristäminen:** pylväskaaviossa y-akselin tulisi alkaa nolasta tai mittarin minimistä, sillä muualta aloittaminen voi liioitella eroja. Viivakaavioissa akselia voi rajata korostamaan pieniä muutoksia, kunhan kokonaiskuva säilyy selkeänä.
- **Värien väärinkäyttö:** liian monet värit tekevät visualisoinnista sekavan. Käytä enintään 5–7 väriä ja pidä värikoodaus loogisena (esim. negatiiviset arvot punaisella, positiiviset vihreällä).
- **Epäsopivan kaaviotyypin valinta:** piirakkakaavio toimii vain, kun vertaillaan 2–6 osaa – useamman osan kohdalla pylväskaavio on selkeämpi. Käytä oikeaa kaaviotyyppiä viestin vahvistamiseksi.
- **Liiallinen informaatio:** kun yhteen visualisointiin ehdetaan liikaa tietoa, se muuttuu vaikealukuseksi. Jaa data useampaan kaavioon tai karsi epäolennainen pois.
- **Puutteellinen otsikointi ja selitteet:** selkeät otsikot, akselien nimet ja selitteet ovat välttämättömiä, jotta katsoja ymmärtää visualisoinnin oikein.

- **Turhat 3D-efektit:** kolmiulotteiset kaaviot saattavat näyttää näyttäviltä, mutta ne usein vääristävät tietoa. Selkeys on tärkeämpää kuin visuaalinen näyttävyys – suosi kaksiulotteisia kaavioita.
- **Datan kontekstin puuttuminen:** ilman vertailukohtia (esim. edellisen vuoden luvut tai toimialan keskiarvo) yksittäinen luku tai trendi voi jäädä merkityksettömäksi. Lisää aina tarvittava taustatieto.
- **Skaalauksen epäjohtonmukaisuus:** jos eri kaavioissa käytetään eri asteikkoja, vertailu voi johtaa harhaan. Pidä skaalat yhtenäisinä, jotta erot ja trendit näkyvät oikeassa mittakaavassa.

Yksi tehokas tapa varmistaa visualisoinnin selkeys on pyytää palautetta ennen sen julkaisua. Kun useampi henkilö tarkastelee esitystä, on helpompi havaita epäselvyydet ja mahdolliset harhaanjohtavat näkökulmat. Lisäksi on tärkeää testata, ymmärtääkö kohderyhmä visualisoinnin oikein – visuaalisesti näyttävä kaavio ei tue päätöksentekoa, jos se ei ole käyttäjälle intuitiivinen.

Visualisoinneissa kannattaa suosia vakiintuneita esitystapoja, sillä tutut kaaviotyypit ja rakenteet helpottavat hahmottamista. Vaikka luovat ratkaisut voivat herättää huomiota, turha monimutkaisuus haittaa tiedon ymmärtämistä. Erityisesti liiketoimintaympäristössä yksinkertaisuus ja selkeys ovat usein tehokkaampia kuin visuaaliset erikoisratkaisut.

On myös olennaista testata, miten visualisointi toimii eri laitteilla ja näytöillä. Kaavio, joka näyttää selkeältä tietokoneen ruudulla, voi muuttua epäselväksi mobiilinäytöllä tai heikkolaatuisella projektorilla. Skaalautuvuuden ja luettavuuden tarkistaminen eri ympäristöissä varmistaa, että data on ymmärrettävää kaikille käyttäjille.

Hyvin toteutettu analyysi ja selkeä visualisointi ovat tehokkaan päätöksenteon kulmakiviä. Ne auttavat hahmottamaan monimutkaisia ilmiöitä ja tukevat organisaatioita hyödyntämään dataa kilpailuetunaan.

## Tehtävä 11

Valitse oikea vaihtoehto.

### **Mikä seuraavista on datan onnistuneen visualisoinnin keskeinen periaate?**

- a) Käytä mahdollisimman paljon värejä, jotta visualisointi olisi näyttävä ja huomiota herättävä.
- b) Valitse aina 3D-kaavio, koska se näyttää modernimmalta ja houkuttelevammalta.
- c) Esitä data selkeästi ja yksinkertaisesti, jotta se on helposti ymmärrettävää ilman tarpeettomia häiriötekijöitä.

- d) Aloita Y-akseli mistä tahansa arvosta, jotta pienet erot saadaan korostettua dramaattisemmin.

*Oikea vastaus: c) Hyvä visualisointi keskittyy yksinkertaisuuteen ja selkeyteen, jotta katsoja voi ymmärtää datan ilman turhia häiriötekijöitä. Liiallinen värien käyttö (a) voi tehdä kaaviosta sekavan, 3D-kaaviot (b) voivat vääristää tietoa, ja y-akselin epäjohdonmukainen skaalaus (d) voi johtaa harhaanjohtaviin tulkintoihin.*

## Tehtävä 12

Pohdi onko väittämä oikein vai väärin.

**Interaktiiviset dashboardit ovat hyödyllisiä, koska ne auttavat käyttäjiä suodattamaan ja tarkastelemaan tietoa eri näkökulmista.**

- a) Oikein
- b) Väärin

*Oikea vastaus: Oikein. Interaktiiviset dashboardit mahdollistavat tietojen syvemmän tarkastelun ja antavat käyttäjille mahdollisuuden tutkia eri muuttujia joustavasti.*

**Piirakkakaavio on paras tapa esittää trendien kehitystä pitkällä aikavälillä.**

- a) Oikein
- b) Väärin

*Oikea vastaus: Väärin. Piirakkakaavio ei sovellu trendien esittämiseen, vaan se on tarkoitettu osien vertailuun kokonaisuudessa. Trendien analysointiin viivakaavio on parempi vaihtoehto.*

**Mitä enemmän värejä visualisoinnissa käytetään, sitä helpommin se on tulkittavissa.**

- a) Oikein
- b) Väärin

*Oikea vastaus: Väärin. Liiallinen värien käyttö voi tehdä visualisoinnista sekavan ja vaikeasti tulkittavan. On suositeltavaa käyttää korkeintaan 5–7 väriä selkeyden säilyttämiseksi.*

**Yksi yleinen virhe datan visualisoinnissa on, että y-akseli ei ala nolasta, mikä voi johtaa harhaanjohtaviin tulkintoihin.**

- a) Oikein

b) Väärin

*Oikea vastaus: Oikein. Jos y-akseli ei ala nollassa, pylväs- tai viivakaaviot voivat antaa harhaanjohtavan kuvan erojen suuruudesta, mikä voi vaikuttaa väärin päätöksentekoon.*

## Yhteenveto

### Luvussa kaksi opimme:

- Datan käsittely alkaa sen keräämisestä, säilyttämisestä ja hallinnasta, ja jokainen vaihe vaikuttaa sen laatuun ja käyttökelpoisuuteen päätöksenteossa.
- Rakenteellinen data on järjestettyä ja helposti analysoitavaa, kun taas rakenteeton data, kuten asiakaspalautteet tai kuvat, vaatii kehittyneempiä analyysimenetelmiä.
- Datan mallintaminen on keskeinen osa tehokasta analytiikkaa, ja faktat sekä dimensioiden määrittely auttavat jäsentämään dataa hyödylliseen muotoon.
- Analytiikka voidaan jakaa neljään tasoon: kuvaileva, diagnosoiva, ennakoiva ja ohjaava, joista jokainen auttaa ymmärtämään ja ennustamaan liiketoiminnan kehitystä eri tavoin.
- Tekoäly tuo analytiikkaan syvyyttä ja tehokkuutta, mahdollistaen suurten datamäärien käsittelyn ja monimutkaisten yhteyksien tunnistamisen.
- Datan eheys, laatu ja saatavuus ovat kriittisiä tekijöitä, sillä ilman luotettavaa dataa analyysit voivat johtaa harhaan ja heikentää päätöksenteon laatua.
- Datan visualisoinnin avulla monimutkaisista tietomassoista voidaan luoda selkeitä ja helposti ymmärrettäviä esitysmuotoja, kuten kaavioita, diagrammeja ja interaktiivisia raportteja.

# Datan voima: Raportointi, mittarit ja eettinen tiedolla johtaminen

Tiedolla johtaminen on yhdistelmä teknologiaa, ihmisten päätöksentekoa ja eettistä vastuuta. Se yhdistää datan ja teknologian liiketoiminnan strategiseen päätöksentekoon. Tässä luvussa keskitymme siihen, miten dataa ja tekoälyä hyödynnetään raportoinnissa, miten suorituskykymittarit tukevat tavoitteiden saavuttamista ja miten eettiset periaatteet varmistavat vastuullisen datan käytön.

## Suorituskykymittarit strategisen johtamisen tukena

Kuten opimme aikaisemmista luvuista, nykypäivän liiketoimintaympäristössä data on yrityksille strateginen resurssi, joka mahdollistaa paitsi päätösten pohjaamisen faktoihin, myös tulevaisuuden ennakkoinnin ja kilpailuedun luomisen. Tiedolla johtaminen on keskeinen työkalu strategisessa päätöksenteossa, sillä se yhdistää datan analyysin, ennusteet ja oivallukset konkreettisiin toimenpiteisiin.

Tiedolla johtaminen olisi kuitenkin vaikeaa ilman oikeaoppista raportointia. Raportointi kokoaa yhteen datan keräämisen, säilytyksen, hallinnan, analysoinnin ja visualisoinnin tulokset selkeäksi ja helposti jaettavaksi kokonaisuudeksi. Raportointi ei ainoastaan tarjoa yhteenvetoa siitä, mitä datasta on opittu, vaan myös auttaa jäsentämään oivallukset päätöksentekoa ja strategista suunnittelua varten.

Miten sitten tiedät miten asiat kehittyvät? Suorituskykymittarit, kuten KPI:t (englanniksi *Key Performance Indicators*, avainsuorituskykymittarit) ja OKR:t (englanniksi *Objectives and Key Results*, tavoitteet ja keskeiset tulokset), ovat kulmakiviä organisaation strategisessa päätöksenteossa. Ne tarjoavat konkreettisia mittareita organisaation suorituskyvyn arviointiin ja resurssien kohdentamiseen.

- **KPI:t** keskittyvät yksittäisiin suorituskykymittareihin, jotka heijastavat organisaation strategisten tavoitteiden toteutumista. Esimerkkejä ovat myyntiluvut, asiakashankinnan kustannus (CAC), asiakastyytyväisyysindeksi (NPS) ja toimitusvarmuus.
- **OKR:t** yhdistävät selkeästi määritellyt tavoitteet (Objectives) ja niiden saavuttamiseen liittyvät konkreettiset tulokset (Key Results). Tämä lähestymistapa luo tiimeille yhtenäisen suunnan ja edistää jatkuvaa parantamista.

KPI:t ja OKR:t ovat suosittuja viitekehyksiä, joiden avulla organisaatiot voivat asettaa ja seurata tavoitteitaan systemaattisesti. Molemmat lähestymistavat tarjoavat numeerisia ja laadullisia mittareita, joiden avulla voidaan arvioida edistymistä ja tehdä tietoon perustuvia päätöksiä parantaen samalla tiimin yhtenäistä suuntaa.

Päätöksenteossa KPI- ja OKR-mittarit toimivat oppaina, jotka auttavat kohdentamaan resursseja ja priorisoimaan tehtäviä. Numeeriset mittarit tarjoavat selkeää ja konkreettista dataa, jonka avulla yritys voi tunnistaa vahvuutensa, paikantaa kehityskohteensa ja kohdentaa resursseja tehokkaasti tavoitteiden saavuttamiseksi.

### Oikeiden KPI- ja OKR-mittareiden valinta käytännön työssä

Suorituskykymittareita valitessa tulee muistaa, että liian monet mittarit tekevät raportoinnista sekavaa, mutta huolella valitut mittarit auttavat keskittymään olennaiseen. Otetaan seuraavaksi esimerkkejä hyvistä KPI- ja OKR-mittareista.

#### Esimerkkejä KPI-mittareista:

- **Markkinoinnissa:** mainoskampanjan konversioprosentti, asiakashankinnan kustannus (englanniksi *CAC*, *Customer Acquisition Cost*), asiakaspoistuma (englanniksi *churn rate*).
- **Myynnissä:** myyntivolyymi, asiakaskohtainen tuotto ja asiakkaan elinkaariarvo (englanniksi *CLV*, *Customer Lifetime Value*), myyntiputken konversioprosentti, joka kertoo kuinka suuri osa tietyistä ihmisistä, kuten verkkosivuston kävijöistä tai mainoksen nähneistä tekee halutun toimenpiteen.
- **Asiakaspalvelussa:** asiakkaiden suosittelemiseksi (englanniksi *NPS*, *Net Promoter Score*), keskimääräinen käsittelyaika (englanniksi *AHT*, *Average Handling Time*), ensiyhteydenoton ratkaisuprosentti (englanniksi *FCR*, *First Call Resolution*).
- **Logistiikassa:** toimitusvarmuus, varastojen kierto nopeus, kuljetuskustannukset per yksikkö.

#### Esimerkkejä OKR-mittareista:

OKR-mittareissa tavoitteet (O) ovat inspiroivia ja selkeitä päämääriä, jotka ohjaavat toimintaa, kun taas avaintulokset (KR) ovat mitattavia ja konkreettisia indikaattoreita, joiden avulla seurataan edistymistä.

Esimerkiksi yritys voi asettaa tavoitteekseen “Parantaa asiakastytyväisyyttä merkittävästi vuoden 2025 aikana.” Tämä tavoite on innostava ja antaa työntekijöille selkeän suunnan. Avaintuloksia voivat olla esimerkiksi:

1. Asiakastytyväisyysmittarin (NPS) nostaminen 70:stä 85:een.
2. Keskeytettyjen tilausten määrän vähentäminen 20 %.
3. Asiakaspalvelupyyntöihin vastaamisen nopeuttaminen 24 tunnista 8 tuntiin.

Nämä avaintulokset ovat konkreettisia, mitattavia ja auttavat työntekijöitä ymmärtämään, mitä heiltä odotetaan. Kun työntekijät näkevät, miten heidän työnsä linkittyy suoraan organisaation laajempiin tavoitteisiin, heidän motivaationsa ja sitoutumisensa lisääntyvät.

OKR-ajattelun avulla organisaatiot voivat ohjata resurssejaan tehokkaammin, seurata edistymistä reaaliaikaisesti ja tehdä tietoon perustuvia päätöksiä. Lisäksi OKR:t edistävät avoimuutta ja tiimien välistä yhteistyötä, koska kaikki työntekijät näkevät, mihin suuntaan organisaatio on menossa ja miten heidän omat ponnistelunsa vaikuttavat kokonaisuuteen. Tämä läpinäkyvyys lisää yhteenkuuluvuuden tunnetta ja parantaa organisaation kykyä sopeutua muutoksiin dynaamisessa liiketoimintaympäristössä.

## Opi lisää

### Muita tapoja mitata kehitystä

- Tasapainotettu tulokortti (Balanced Scorecard): Laaja-alainen lähestymistapa, joka yhdistää taloudelliset ja ei-taloudelliset mittarit neljään näkökulmaan: taloudellinen, asiakasnäkökulma, sisäiset prosessit ja oppiminen ja kasvu.
- Benchmarking: Organisaation suorituskyvyn vertaaminen alan parhaisiin käytäntöihin tai kilpailijoihin, jotta voidaan tunnistaa parannusmahdollisuuksia.
- Agile-mittarit: Mittareita, kuten tiimin nopeus ja tehtävien valmistumisprosentti, käytetään erityisesti ketterissä kehitysprosesseissa.
- Laadullinen palaute: Asiakaskyselyt, työntekijöiden palautteet ja fokusryhmähaastattelut, jotka tarjoavat syvempää ymmärrystä kehityskohteista ja vahvuuksista.
- Data-analytiikka ja ennakoiva analytiikka: Suorituskyvyn ja tulevien trendien arviointi hyödyntämällä suuria tietomääriä ja edistyneitä analytiikkatyökaluja.

## Tehtävä 13

Valitse oikea vastaus.

### Miten suorituskykymitarit tukevat tiedolla johtamista?

- a) Ne auttavat organisaatioita seuraamaan tavoitteidensa toteutumista.
- b) Ne keskittyvät vain henkilöstön hyvinvoinnin mittaamiseen.

- c) Ne auttavat johtajia miettimään omia tavoitteitaan.
- d) Ne mittaavat vain taloudellisia tuloksia ilman vaikutusta muihin toimintoihin.

*Oikea vastaus: a) Suorituskykymittarit, kuten KPI:t (Key Performance Indicators), auttavat organisaatioita arvioimaan, kuinka hyvin ne etenevät strategisten tavoitteidensa mukaisesti.*

### **Mitkä ovat tiedolla johtamisen kolme keskeistä periaatetta strategisessa päätöksenteossa?**

- a) Intuitio, kokemus ja mielipiteet
- b) Hierarkkinen päätöksenteko ilman analytiikkaa
- c) Nopeus, riskinotto ja kokeilut
- d) Objektivisuus, läpinäkyvyys ja ennakoitavuus

*Oikea vastaus: d) Tiedolla johtaminen perustuu objektiiviseen ja mitattavaan dataan, läpinäkyvyyteen organisaation sisällä ja kykyyn ennakoida tulevia tapahtumia tiedon pohjalta.*

## Reaaliaikainen data kilpailuedun lähteenä

Tekoäly on mullistanut raportoinnin roolin organisaatioissa, tehden siitä dynaamisen ja reaaliaikaisen työkalun, joka tukee tavoitteiden seuraamista, kehityssuuntien tunnistamista ja päätöksenteon parantamista kaikilla organisaation tasoilla. Perinteisten staattisten kuukausi- tai kvartaaliraporttien sijaan tekoälyn mahdollistamat reaaliaikainen data ja ennakoiva analytiikka tarjoavat nyt yrityksille mahdollisuuden nopeaan reagointiin ja strategiseen suunnitteluun.

Katsotaan seuraavaksi muutamia esimerkkejä tekoälyn tukemasta raportoinnista:

- **Markkinointi:** tekoäly analysoi kampanjan reaaliaikaisia tuloksia ja ehdottaa, miten budjettia voidaan kohdentaa tehokkaammin. Esimerkiksi verkkokauppa voi seurata mainoskampanjan klikkaus- ja ostokonversioita reaaliajassa ja siirtää panostuksia parhaiten toimiviin kanaviin.
- **Myynti:** myyntitiimit voivat hyödyntää tekoälypohjaisia suosituksia tunnistaakseen potentiaaliset asiakkaat ja tarjota heille sopivimpia tuotteita. Tekoäly voi havaita, että asiakas, joka on ostanut tietyn tuotteen, on todennäköisesti kiinnostunut lisäpalvelusta, kuten huoltosopimuksesta.
- **HR ja henkilöstöhallinto:** tekoäly voi tunnistaa trendejä, kuten työntekijöiden tyytyväisyysmuutoksia, ja ehdottaa toimenpiteitä vaihtuvuuden ehkäisemiseksi. Esimerkiksi suuryritys voi analysoida tyytyväisyyskyselyiden tuloksia ja havaita, että etätyömahdollisuuksien lisääminen parantaa työntekijöiden sitoutumista.

- **Logistiikka ja toimitusketju:** kuljetusten optimointi perustuu reaaliaikaisiin toimitus- ja sääolosuhteisiin. Tekoäly voi ehdottaa vaihtoehtoisia reittejä ruuhkien tai sääolosuhteiden perusteella, mikä säästää sekä aikaa että kustannuksia.

Reaaliaikainen raportointi tuo selkeitä etuja: se mahdollistaa nopeamman reagoinnin, vähentää manuaalista työtä ja auttaa kohdistamaan resursseja paremmin.

Miten tekoälyä sitten voidaan hyödyntää edellisessä kappaleessa esittelemiemme OKR ja KPI-suorituskykymittareiden kehityksessä? Tekoäly voi analysoida OKR ja KPI-mittareiden kehitystä paitsi nykytilanteen perusteella, myös ennakoimalla tulevia trendejä historiallisten ja reaaliaikaisten tietojen pohjalta. Se ei pelkästään tunnista mahdollisia ongelmia, kuten asiakaspoistuman kasvua, vaan myös ehdottaa ennakoivia ja kohdennettuja toimenpiteitä niiden ratkaisemiseksi. Esimerkiksi, jos asiakaspoistuma näyttää kasvavan tietyssä asiakassegmentissä, tekoäly voi analysoida syitä, kuten hintaherkkyyttä tai tyytymättömyyttä tiettyihin palveluihin, ja ehdottaa konkreettisia toimia, kuten personoituja tarjouksia, kohdennettuja kampanjoita tai asiakaspalvelun kehittämistä. Lisäksi tekoäly voi simuloida erilaisten strategioiden vaikutuksia ennen niiden käyttöönottoa, mikä auttaa yritystä valitsemaan tehokkaimmat toimenpiteet ongelman ratkaisemiseksi.

## Opi lisää

### Tekoälyavusteinen raportointi ja automaattiset mittaristot (englanniksi *dashboard*)

Tekoälypohjaiset dashboardit ovat nykyaikaisen raportoinnin ydin. Ne yhdistävät tekoälyn, datan analytiikan ja visualisoinnin, tarjoten reaaliaikaista tietoa liiketoiminnan keskeisistä mittareista. Tämä mahdollistaa nopeamman reagoinnin, tehokkaamman päätöksenteon ja resurssien paremman kohdentamisen.

### Mitä hyötyä tekoälypohjaisista dashboardeista on?

- **Reaaliaikainen seuranta**  
Näe tärkeimmät liiketoimintamittarit yhdellä silmäyksellä, kuten myynti, asiakastytyväisyys ja markkinatrendit.
- **Poikkeamien havaitseminen**  
Tekoäly tunnistaa ja hälyttää heti, kun jokin mittari poikkeaa normaalista.
- **Toimintasuositukset**  
Tekoäly voi analysoida dataa ja ehdottaa keinoja esimerkiksi markkinoinnin optimointiin tai asiakaspalvelun parantamiseen.
- **Automatisointi ja ennustaminen**  
Dashboardit voivat paitsi raportoida nykytilanteen, myös ennakoida tulevia kehityssuuntia.

## Tehtävä 14

Valitse oikea vastaus.

### **Mikä seuraavista on reaaliaikaisen datan ja tekoälyn tuoma keskeinen hyöty organisaatioille?**

- a) Se korvaa kokonaan ihmisten tekemän päätöksenteon, jolloin yritykset voivat toimia täysin automaattisesti.
- b) Se mahdollistaa nopeamman reagoinnin muuttuviin tilanteisiin ja resurssien tehokkaamman kohdentamisen.
- c) Se vähentää yritysten tarvetta seurata suorituskykymittareita (OKR ja KPI), koska tekoäly hoitaa kaiken analyysin itsenäisesti.
- d) Se tekee perinteiset raportit ja analytiikkatyökalut tarpeettomiksi, sillä organisaatioille riittää tekoälyavusteinen dashboard.

*Oikea vastaus: b) Reaaliaikainen data ja tekoäly tukevat päätöksentekoa tarjoamalla ajantasaista tietoa, tunnistamalla kehityssuuntia ja ennakoimalla mahdollisia haasteita. Tämä mahdollistaa yrityksille nopeamman reagoinnin ja resurssien kohdentamisen esimerkiksi markkinoinnissa, myynnissä ja toimitusketjun hallinnassa. Vaihtoehdot a, c ja d ovat harhaanjohtavia, koska tekoäly ei täysin korvaa ihmisten päätöksentekoa, suorituskykymittareiden seuranta tai perinteisiä analyysimenetelmiä, vaan tukee niitä tehokkaammin.*

## Kuinka johtaa tiedolla eettisesti?

Tiedolla johtaminen ei ole vain teknologiaa ja data-analytiikkaa, vaan ennen kaikkea ihmisten välistä yhteistyötä ja ajattelutavan muutosta. Datalukutaito, kriittinen ajattelu ja eettinen harkinta ovat keskeisiä taitoja tiedon hyödyntämisessä – ilman kykyä ymmärtää ja soveltaa dataa liiketoiminnassa, teknologia voi jäädä hyödyntämättömäksi tai johtaa tahattomiin haitallisiin seurauksiin. Organisaatioiden on panostettava työntekijöidensä osaamisen kehittämiseen, jotta dataa osataan käyttää paitsi tehokkuuden lisäämiseen myös eettisesti ja vastuullisesti.

Eettinen tiedolla johtaminen varmistaa, että organisaation päätökset perustuvat luotettavaan ja oikeudenmukaiseen dataan, eivätkä johda epäoikeudenmukaisiin seurauksiin. Esimerkiksi, jos rekrytointiprosessissa käytetään tekoälyä arvioimaan työnhakijoita, mutta algoritmi on koulutettu puolueellisella datalla, se voi suosia tai syrjiä tiettyjä hakijaryhmiä. Tällaisten vinoumien tunnistaminen ja korjaaminen on keskeinen osa eettistä tiedolla johtamista.

Tiedolla johtamisen avulla yritykset voivat paitsi optimoida päätöksentekoa ja kehittää liiketoimintaa, myös luoda ratkaisuja, jotka tukevat kestäväää kehitystä ja yhteiskunnallista vastuullisuutta. Esimerkiksi tekoälyn ja datan avulla voidaan optimoida toimitusketjuja, vähentää turhia resursseja ja pienentää ympäristövaikutuksia, mikä tukee vastuullista yritystoimintaa. Samalla suorituskykymittarit mahdollistavat sen, että toimintaa voidaan mitata datalähtöisesti, asettaa selkeitä tavoitteita ja seurata niiden toteutumista.

Yritykset, jotka osaavat yhdistää datan ja teknologian eettisiin toimintaperiaatteisiin, eivät ainoastaan paranna kilpailukykyään, vaan myös vahvistavat asiakkaiden, sidosryhmien ja työntekijöiden luottamusta. Läpinäkyvä ja vastuullinen datan käyttö voi toimia brändin kilpailuetuna, erityisesti tilanteissa, joissa asiakkaat ja yhteistyökumppanit arvostavat eettisiä ja vastuullisia toimintatapoja.

## Opi lisää

### Eettisen tiedolla johtamisen keskeiset periaatteet

Eettinen tiedolla johtaminen varmistaa, että dataa kerätään, analysoidaan ja hyödynnetään oikeudenmukaisesti sekä yksilöiden oikeudet huomioiden.

Keskeiset periaatteet eettisessä tiedolla johtamisessa:

- **Läpinäkyvyys**  
Käyttäjien tulee ymmärtää, miten heidän tietojansa kerätään ja käytetään. Selkeä viestintä ja tietosuojaselosteet lisäävät luottamusta.
- **Anonymiteetti**  
Datan anonymisointi suojaa yksityisyyttä ja mahdollistaa turvallisen analytiikan ilman tunnistettavia henkilötietoja.
- **Datan minimointi**  
Kerätään vain välttämätön tieto, mikä vähentää riskejä ja tehostaa datan hyödyntämistä.
- **Vastuullinen käyttö**  
Algoritmien ja mallien tulee olla syrjimättömiä ja oikeudenmukaisia. Tekoälyratkaisuja on seurattava vinoumien tunnistamiseksi.
- **Tietoturva**  
Henkilötiedot on suojattava salauksella, monivaiheisella tunnistautumisella ja käyttöoikeuksien hallinnalla.
- **Henkilöstön koulutus**  
Työntekijöiden tulee tuntea tietosuoja- ja tietoturvakäytännöt, kuten GDPR:n vaatimukset.

## Tehtävä 15

Valitse oikea vastaus.

### **Mikä on eettisen tiedolla johtamisen tavoite?**

- a) Datan kerääminen mahdollisimman laajasti ilman rajoituksia
- b) Käyttää dataa vastuullisesti, oikeudenmukaisesti ja läpinäkyvästi päätöksenteossa
- c) Tehostaa päätöksentekoa ilman, että eettisiä riskejä arvioidaan
- d) Korvata ihmisten päätöksenteko täysin tekoälyllä

*Oikea vastaus: b) Eettinen tiedolla johtaminen varmistaa, että dataa käytetään oikeudenmukaisesti ja yksilöiden oikeuksia kunnioittaen. Se ei tarkoita datan keräämistä rajattomasti tai ihmisten päätöksenteon korvaamista tekoälyllä.*

## Yhteenveto

### **Luvussa kolme opimme:**

- Tiedolla johtaminen on strategisen päätöksenteon keskeinen osa, ja sen avulla organisaatiot voivat tehdä tietoon perustuvia, objektiivisia ja ennakoitavia päätöksiä.
- Datan hyödyntäminen päätöksenteossa parantaa organisaation kilpailukykyä, koska se mahdollistaa paremman riskienhallinnan, resurssien tehokkaamman kohdentamisen ja pitkän aikavälin tavoitteiden saavuttamisen.
- Suorituskykymittarit kuten KPI:t ja OKR:t ovat tiedolla johtamisen keskeinen työkalu, ja niiden avulla voidaan seurata päätösten vaikutuksia sekä arvioida organisaation menestystä.
- Tiedolla johtaminen ei ole vain tekninen prosessi, vaan se vaatii organisaatiokulttuurin muutosta, jossa päätöksenteko perustuu dataan, analyytiikkaan ja jatkuvaan oppimiseen intuitiivisten päätösten sijaan.
- Eettisyys ja vastuullisuus ovat olennaisia tekijöitä tiedolla johtamisessa, sillä datan käyttöön liittyy yksityisyydensuojaan, vinoumien hallintaan ja läpinäkyvyyteen liittyviä kysymyksiä, jotka vaikuttavat organisaation maineeseen ja kestävään kasvuun.

# Datan tuomat mahdollisuudet liiketoiminnan tulevaisuudessa

Data ei ole enää pelkästään operatiivinen resurssi, vaan strateginen voimavara, joka määrittää yritysten kilpailukyvyn ja innovaatiokyvyn. Tämä luku syventyy siihen, miten dataa voidaan kaupallistaa, millaisia liiketoimintamalleja sen ympärille voi rakentaa ja kuinka datan hyödyntäminen muokkaa tulevaisuuden liiketoimintaympäristöä.

## Datan kaupallistaminen

Datan kaupallistaminen eli monetisointi tarkoittaa prosessia, jossa dataa hyödynnetään taloudellisen arvon luomiseksi. Tämä voi tapahtua monella tavalla, kuten datan suoralla myynnillä, palveluihin lisäarvon tuottamisella tai kokonaan uusien liiketoimintamallien kehittämisellä. Kuten aikaisemmista kappaleista opimme, data ei ole itsessään arvokasta – sen todellinen arvo syntyy, kun se jalostetaan oivalluiksi, päätöksiksi ja uusiksi mahdollisuuksiksi.

### Datan kaupallistamisen tavat

#### 1. Datan lisensointi

- **Määritelmä:** dataa tarjotaan toisille yrityksille käytettäväksi tietyin ehdoin ja rajoitetuksi ajaksi. Lisensoinnin avulla yritys säilyttää datan omistajuuden, mutta mahdollistaa sen hyödyntämisen kumppanien toimesta.
- **Käytännön esimerkki:** asiakkaiden ostokäyttäytymistiedot voivat olla arvokkaita markkinatutkimusyrityksille, jotka analysoivat kuluttajatrendejä.

#### 2. Raakadatan myynti

- **Määritelmä:** jalostamatonta dataa myydään suoraan eteenpäin toisille yrityksille, jotka voivat jalostaa tai analysoida sitä tarpeisiinsa.
- **Käytännön esimerkki:** teknologia- ja sensorirytykset myyvät raakadataa esimerkiksi liikenneanalyysiin tai sään ennustamiseen.

### 3. Analytiikkapalvelut

- **Määritelmä:** organisaatiot voivat kaupallistaa analyysituloksia, ennusteita tai visualisointeja asiakkaiden liiketoiminnan tueksi.
- **Käytännön esimerkki:** yritys voi tarjota ennustavaa analytiikkaa myyntitietojen perusteella, jotta asiakkaat voivat optimoida varastotasojaan.

### 4. Ekosysteemien luominen

- **Määritelmä:** rakennettu alusta, jossa eri organisaatiot voivat jakaa ja ostaa dataa keskenään.
- **Käytännön esimerkki:** datamarkkinapaikat, joissa dataa tarjoavat yritykset voivat yhdistää voimansa muiden toimijoiden kanssa luodakseen uusia palveluita tai tuotteita.

Kuten yllä olevista esimerkeistä huomaamme, datan monetisointi voi olla suoraa, kuten anonymisoidun datan myyntiä markkinatutkimuksiin, tai epäsuoraa, esimerkiksi parantamalla prosessien tehokkuutta tai optimoimalla asiakaskokemusta. Tutustutaan seuraavaksi datan monetisointiin käytännön case-esimerkin avulla.

#### Case

#### Uber, datan hyödyntämisen mestari

Uber on yksi tunnetuimmista esimerkeistä yrityksestä, joka on rakentanut liiketoimintansa datan hyödyntämisen ympärille. Yrityksen menestys perustuu sen kykyyn kerätä, analysoida ja hyödyntää valtavia määriä dataa reaaliajassa. Tämä mahdollistaa sekä paremman asiakaskokemuksen että operatiivisen tehokkuuden.

#### Miten Uber hyödyntää dataa?

1. Reaaliaikainen sijaintidata ja reittioptimointi  
Uber käyttää GPS-sijaintitietoja sekä kuljettajilta että matkustajilta reittien suunnitteluun. Tämä mahdollistaa nopeimman ja kustannustehokkaimman reitin valinnan. Kuljettajat näkevät parhaan reitin kohteeseen, ja matkustajat hyötyvät lyhyemmistä odotusajoista ja kustannustehokkaammista matkoista.
2. Dynaaminen hinnoittelu  
Uberin algoritmit analysoivat reaaliajassa kysynnän ja tarjonnan tasapainoa. Jos tietyllä alueella on paljon matkustajia ja vain vähän kuljettajia, hinta nousee kysyntäpiikin hallitsemiseksi. Tämä niin sanottu "surge pricing" maksimoi tuloja ja kannustaa kuljettajia siirtymään kysyntäalueille.
3. Ennakoiva analytiikka ja kysynnän ennustaminen  
Uber hyödyntää historiallista dataa ja koneoppimista ennustaakseen, missä ja milloin kuljettajia tarvitaan eniten. Esimerkiksi suurten tapahtumien aikana Uber voi sijoittaa

kuljettajia strategisesti, mikä vähentää odotusaikoja ja kasvattaa tuloja.

4. Asiakaskokemuksen parantaminen

Uber seuraa matkustajien ja kuljettajien palautetta sekä matkojen suorituskykyä. Datan avulla se voi tehdä jatkuvia parannuksia, kuten optimoida sovelluksen käyttöliittymää tai tunnistaa ongelmatilanteet, joissa tarvitaan asiakaspalvelua.

5. Kuljettajien suorituskyvyn hallinta

Uber analysoi kuljettajien arvioita, ajokäyttäytymistä ja aktiivisuutta varmistaakseen palvelun laadun. Tämä auttaa myös palkitsemaan aktiivisia ja korkeasti arvioituja kuljettajia.

### Datan kaupallistaminen ja liiketoimintamalli

Uberin liiketoimintamalli perustuu alustan tarjoamiseen, joka yhdistää matkustajat ja kuljettajat. Vaikka Uber ei suoraan myy keräämäänsä dataa, se käyttää sitä tehokkaasti uusien palvelujen kehittämiseen, operatiivisten kustannusten vähentämiseen sekä mainonnan kohdentamiseen. Lisäksi data toimii pohjana laajentumisstrategialle: uusille markkinoille siirtyminen on nopeampaa, kun aiemmilta markkinoilta kerätty tieto ohjaa toimintaa.

Uberin liiketoimintamallin ydin:

- Datan käyttö arvonluonnissa: optimointi, hinnoittelu, asiakaskokemus.
- Alustatalous: Uber toimii välittäjänä matkustajien ja kuljettajien välillä, ja se ottaa osan matkan hinnasta palvelumaksuna.
- Mittakaavaetu (Englanniksi *Economies of scale*): mitä enemmän dataa Uber kerää, sitä tarkemmin sen algoritmit toimivat, mikä parantaa palvelua entisestään.

### Mitä voimme oppia Uberin mallista?

Uberin esimerkki korostaa, kuinka tärkeää on:

1. Reaaliaikainen datan analysointi: nopea päätöksenteko parantaa asiakastyytyväisyyttä ja tehokkuutta.
2. Datan strateginen hyödyntäminen: ei pelkästään operatiivisiin tarpeisiin, vaan myös innovaatioiden ja uusien markkinoiden avauksessa.
3. Asiakaskokemuksen jatkuva parantaminen: data toimii pohjana palautteiden käsittelylle ja palvelujen kehittämiselle.

### Datan monetisointi Suomessa

Suomen vahvuudet – korkea teknologiaosaaminen, toimiva infrastruktuuri ja tiukka tietosuojalainsäädäntö – tekevät maastamme ihanteellisen ympäristön datan monetisoinnille. Monetisointia hyödynnetään Suomessa esimerkiksi energiasektorilla ja terveysteknologiassa.

Suomessa yhdistyvät vahva teknologiaosaaminen, tiukka tietosuojalainsäädäntö ja digitalisoitunut yhteiskunta, mikä luo erinomaiset edellytykset datatalouden kehittämiseksi. Suomen yritykset ja

tutkimuslaitokset ovat huipputasoa data-analytiikan osaamisessa, mikä tukee sekä innovaatioita että kilpailukykyä. Lisäksi tiukka GDPR-säädösten, eli Euroopan tietosuoja-lainsäädännön noudattaminen luo luottamusta ja mahdollistaa datan turvallisen kaupallistamisen.

Suomen digitalisoitunut yhteiskunta tarjoaa vahvan infrastruktuurin datan hyödyntämiseen. Älykkäät kaupungit, digitaaliset palvelut ja laaja tietoverkko mahdollistavat datan ja sen myötä tekoälyn innovatiivisen käytön eri sektoreilla. Tämä yhdistelmä tekee Suomesta edelläkävijän tietojohdantamisen ja datatalouden edistämisessä.

## Case

### Nightingale Health – data terveyden edistäjänä

Nightingale Health on suomalainen terveydenhuoltoalan yritys, joka hyödyntää dataa ennakoivan terveydenhuollon kehittämiseen. Yhtiö tunnetaan erityisesti edistyneestä verianalyysitekniikastaan, joka tarjoaa syvällistä tietoa yksilön terveydentilasta ja mahdollistaa sairauksien riskien tunnistamisen ennen niiden puhkeamista.

### Miten Nightingale Health hyödyntää dataa?

1. Laaja verianalyysidata  
Nightingale kerää verinäytteistä kattavaa dataa, kuten biomarkkereita, jotka paljastavat riskitekijöitä kroonisiin sairauksiin, kuten sydän- ja verisuonitauteihin ja tyypin 2 diabetekseen. Tämä data yhdistetään biopankkiaineistoihin ja muihin terveystietoihin syvällisen analyysin mahdollistamiseksi.
2. Ennakoiva analytiikka  
Yhtiön analyysityökalut hyödyntävät koneoppimista ja tekoälyä riskien arviointiin. Esimerkiksi Nightingalen järjestelmät voivat tunnistaa yksilöt, joilla on kohonnut riski sairastua tiettyihin sairauksiin, ja tarjota kohdennettuja suosituksia elintapojen muuttamiseksi.
3. Yksilölliset terveystiedot  
Nightingale muuntaa monimutkaiset biomarkeritiedot selkeiksi ja ymmärrettäviksi raporteiksi, joita asiakkaat ja terveydenhuollon ammattilaiset voivat käyttää. Raportit tarjoavat konkreettisia tietoja, kuten kolesterolitasot, sokeriaineenvaihdunnan tilan ja tulehdustekijät.
4. Tutkimus ja innovaatio  
Nightingale tekee yhteistyötä biopankkien ja tutkimusorganisaatioiden kanssa luodakseen laajoja tietokantoja, jotka auttavat ymmärtämään sairauksien kehittymistä ja niiden ennaltaehkäisyä. Tämä tutkimustieto voi johtaa uusiin hoitomuotoihin ja terveysinnovaatioihin.
5. Datapohjaiset palvelut yrityksille  
Nightingale tarjoaa datapalveluja vakuutusyhtiöille, terveydenhuollon toimijoille ja hyvinvointialoille, jotka voivat käyttää analytiikkaa asiakasohjelmiensa ja palveluidensa parantamiseen. Näin yritys kaupallistaa dataansa epäsuorasti tarjoamalla lisäarvoa kumppaneilleen.

## Datan kaupallistaminen ja liiketoimintamalli

Nightingale Healthin liiketoimintamalli perustuu datan hyödyntämiseen terveyden optimoinnissa. Yritys tarjoaa kuluttajille ja yrityksille palveluja, jotka perustuvat sen kehittämään verianalyysiteknologiaan. Tuloja syntyy sekä suorista asiakasmaksuista että yhteistyösopimuksista yritysten ja organisaatioiden kanssa.

Nightingalen liiketoimintamallin avainkohdat:

- Kuluttajapalvelut: Asiakkaat voivat tilata verianalyysin ja saada yksilöllisiä terveystietoja.
- Yritysyhteistyö: Vakuutusyhtiöt ja terveydenhuollon toimijat hyödyntävät Nightingalen analytiikkaa palvelujensa kehittämiseen.
- Tutkimuslisenssit: Yhtiö tarjoaa biopankkiaineistoihin perustuvaa dataa tutkimusorganisaatioille innovaatioiden tukemiseksi.

## Mitä voimme oppia Nightingale Healthin mallista?

Nightingale Health tarjoaa konkreettisia esimerkkejä siitä, kuinka data voi olla sekä yhteiskunnallisesti merkittävä että kaupallisesti arvokas. Tärkeimmät opit:

1. Ennakoiva terveydenhuolto: Datan avulla voidaan siirtyä reaktiivisesta hoidosta ennaltaehkäisevään terveydenhuoltoon, mikä parantaa ihmisten elämänlaatua ja vähentää kustannuksia.
2. Yhteistyö biopankkien ja tutkimuslaitosten kanssa: Suuret tietokannat mahdollistavat uusien oivallusten löytämisen ja innovaatioiden kehittämisen.
3. Monipuolinen liiketoimintamalli: Nightingale yhdistää suoran asiakaspalvelun, yritysyhteistyön ja tutkimuksen, mikä tekee liiketoiminnasta monipuolista ja vakaata.

## Dataan liittyvät liiketoimintamallit

Datapohjaisten liiketoimintamallien suosio on kasvanut merkittävästi, sillä ne mahdollistavat skaalautuvuuden, joustavuuden ja tasaisen tulovirran luomisen. Data-as-a-Serviceen (DaaS) on yksi merkittävimmistä datan kaupallistamisen muodoista. DaaS tarjoaa organisaatioille keinon hyödyntää dataa helposti saatavilla olevana palveluna, mikä mahdollistaa datan jakamisen, analysoinnin ja arvon tuottamisen ilman merkittäviä omia infrastruktuuri-investointeja.

**Data-as-a-Service** on liiketoimintamalli, jossa organisaatiot tarjoavat dataa palveluna muiden yritysten, viranomaisten tai yksilöiden käytettäväksi. DaaS-palvelut mahdollistavat datan reaaliaikaisen jakelun rajapintojen (API) kautta, jolloin käyttäjät voivat helposti integroida ja hyödyntää dataa omissa järjestelmissään ja prosesseissaan.

Nämä tarjoavat reaaliaikaista dataa esimerkiksi maanviljelijöille ja logistiikkayrityksille sekä esimerkiksi energiantuotannon optimointiin, laajentaen perinteisiä ansaintamalleja ja vahvistaen datan ympärille rakentuvia liiketoimintaekosysteemejä.

## Case

### Data-as-a-Service

**Esimerkkiyritys:** Ilmatieteenlaitos

**Ilmatieteenlaitoksen säädata** on erinomainen esimerkki DaaS-palvelusta Suomessa. Ilmatieteenlaitos tarjoaa monipuolista säähän liittyvää dataa eri käyttäjäryhmille, kuten yrityksille, tutkijoille ja viranomaisille. Tämä data on saatavilla rajapintojen (API) kautta, mikä mahdollistaa sen integroimisen erilaisiin sovelluksiin ja järjestelmiin.

### DaaS-palvelun tarjoamat hyödyt:

- 1. Reaaliaikainen säädata**
  - Lämpötila, tuulen nopeus, sademäärä, ilmanpaine, näkyvyys
  - Käytetään esimerkiksi logistiikan ja rakennusalan suunnittelussa.
- 2. Sääennusteet**
  - Lyhyen ja pitkän aikavälin ennusteet
  - Soveltuvat esimerkiksi tapahtumajärjestäjille ja maanviljelijöille.
- 3. Ilmastodata**
  - Pitkän aikavälin keskiarvot ja trendit
  - Hyödyllistä ympäristötutkimuksessa ja energiasuunnittelussa.
- 4. Erikoisdatasetit**
  - Salamatulkatiedot, jäättilanteen seuranta, ilmanlaatu.
  - Käytetään mm. lentoliikenteessä ja turvallisuussuunnittelussa.

Säädatan käyttökohteita:

- **Logistiikka ja kuljetus:** optimoidaan reittejä ja aikatauluja sääolosuhteiden mukaan.
- **Rakennusala:** sääolosuhteiden huomiointi rakennusprojekteissa, kuten betonin valussa tai nosturitöissä.
- **Maatalous:** päätökset kastelusta, kylvöstä tai sadonkorjuusta perustuvat sääennusteisiin.
- **Vakuutusyhtiöt:** vahinkotilanteiden ennakointi ja riskienhallinta.

Ilmatieteen laitoksen säädata on saatavilla useissa muodoissa, kuten reaaliaikaisena virtaavana datana, ladattavina tiedostoina tai API-rajapintojen kautta. Tämä tekee siitä joustavan työkalun monille eri toimialoille.

### Taloudellinen malli:

Ilmatieteenlaitoksen taloudellinen malli yhdistää julkisen rahoituksen ja maksulliset palvelut.

Julkinen rahoitus kattaa kansalaisille maksuttomat peruspalvelut, kuten sääennusteet ja viranomaisten tarvitsemat turvallisuuskriittiset palvelut. Lisäksi laitos myy tarkkaa ja räätälöityä dataa yrityksille, kuten logistiikka- ja energiasektorille, sekä tarjoaa konsultointia ja erikoispalveluita, kuten salamatutkatietoja ja jääkarttoja. Tämä hybridimalli mahdollistaa laajan palvelutarjonnan, innovaatioiden kehittämisen ja datan korkean laadun ylläpitämisen. Samalla se varmistaa, että peruspalvelut ovat kaikkien saatavilla ilman erillistä maksua.

## Tehtävä 16

Valitse oikea vastaus.

### **Mikä seuraavista erottaa datan lisensoinnin ja raakadatan myynnin toisistaan?**

- a) Lisensoinnissa dataa myydään eteenpäin ilman rajoituksia, kun taas raakadatan myynnissä data jää myyjälle
- b) Lisensoinnissa yritys säilyttää datan omistajuuden ja antaa sen käyttöön tietyin ehdoin, kun taas raakadatan myynnissä data siirtyy ostajan omistukseen.
- c) Lisensointi koskee vain anonymisoitua dataa, kun taas raakadata sisältää aina henkilökohtaisia tietoja
- d) Raakadatan myynti on sallittua vain teknologia-alalla, kun taas lisensointi on yleinen käytäntö kaikilla aloilla.

*Oikea vastaus: b) Lisensoinnissa yritys säilyttää datan omistajuuden ja antaa sen käyttöön tietyin ehdoin, kun taas raakadatan myynnissä data siirtyy ostajan omistukseen.*

### **Mikä seuraavista EI ole datan kaupallistamisen muoto?**

- a) Datan lisensointi
- b) Raakadatan myynti
- c) Datan käyttö asiakaspalvelun kehittämiseen
- d) Analytiikkapalvelut
- e) Ekosysteemien luominen

*Oikea vastaus: c) Datan käyttö asiakaspalvelun kehittämiseen on tärkeää, mutta se ei ole kaupallistamisen muoto, vaan enemmänkin liiketoiminnan optimointia.*

### **Mitä tarkoittaa Data-as-a-Service (DaaS)?**

- a) Datan myyminen suoraan asiakkaille staattisina tiedostoina ilman jatkohyödyntämismahdollisuuksia.
- b) Datan tarjoaminen palveluna, jossa asiakkaat voivat käyttää ja analysoida tietoa reaaliaikaisesti rajapintojen (API) kautta.
- c) Yrityksen sisäinen datanhallintajärjestelmä, joka auttaa päätöksenteossa mutta ei ole ulkopuolisten saatavilla.
- d) Organisaation oma datakeskus, jossa tietoa säilytetään, mutta jota ei jaeta muiden toimijoiden kanssa.

*Oikea vastaus: b) DaaS tarkoittaa datan tarjoamista palveluna, jonka asiakkaat voivat hyödyntää ilman, että heidän täytyy itse investoida laajoihin infrastruktuureihin tai tietojen käsittelyyn. Se mahdollistaa datan helpon integroinnin eri järjestelmiin ja sovelluksiin reaaliaikaisesti API-rajapintojen kautta.*

## Tulevaisuuden liiketoiminta rakentuu datan varaan

Datan rooli liiketoiminnassa on kehittynyt pelkästä operatiivisesta tukifunktiosta strategiseksi kilpailutekijäksi. Tiedon tuottama lisäarvo riippuu siitä, kuinka hyvin organisaatio hallitsee dataansa ja rakentaa sen pohjalta skaalautuvia ratkaisuja. Tulevaisuuden liiketoiminta on datapohjaista ja ihmiskeskeistä, jossa teknologiaa hyödynnetään vastuullisesti ja eettisesti.

Tulevaisuuden menestyjät ymmärtävät, että datalukutaito ja tekoälylukutaito ovat keskeisiä osaamisalueita. World Economic Forumin *Tulevaisuuden työ 2025* -raportin mukaan jopa 39 prosenttia nykyisistä taidoista vanhentuu tai vaatii päivitystä vuoteen 2030 mennessä. Tämä korostaa jatkuvan oppimisen ja osaamisen kehittämisen tarvetta. Yritysten on yhdistettävä tekninen ja inhimillinen näkemys rakentaakseen oikeudenmukaisia, eettisiä ja liiketoiminnallisesti kestäviä ratkaisuja – ja tähän tarvitaan laadukasta dataa.

### Käsite

#### **Datalukutaito ja tekoälylukutaito**

**Datalukutaito** tarkoittaa kykyä ymmärtää, tulkita ja hyödyntää dataa päätöksenteossa. Se sisältää esimerkiksi datan keräämisen periaatteet, analysointimenetelmät ja kyvyn arvioida datan luotettavuutta. Datalukutaitoinen henkilö osaa tunnistaa harhaanjohtavia tilastoja ja ymmärtää, miten data vaikuttaa liiketoiminnallisiin tai yhteiskunnallisiin päätöksiin.

**Tekoälylukutaito** tarkoittaa ymmärrystä siitä, miten tekoäly toimii, miten sitä voidaan hyödyntää

ja mitkä ovat sen rajoitteet ja eettiset vaikutukset. Se ei edellytä teknistä osaamista tai ohjelmointitaitoja, vaan ennen kaikkea kykyä arvioida tekoälyn tuottaman tiedon luotettavuutta ja sen käyttöä eri tilanteissa.

## Keskeiset taidot tulevaisuuden liiketoiminnassa

Datapohjaisen mutta ihmiskeskeisen liiketoiminnan menestyksen kannalta olennaisia taitoja ovat:

- **Datalukutaito ja kontekstuaalinen ymmärrys:** kyky arvioida kriittisesti datasta saatuja tuloksia ja välttää harhaanjohtavia johtopäätöksiä.
- **Dataviestintätaidot:** kyky esittää dataan perustuvia havaintoja selkeästi ja ymmärrettävästi eri sidosryhmille, kuten asiakkaille, tiimeille ja johdolle. Hyvä viestintä varmistaa, että analyysin tulokset muuttuvat käytännön liiketoimintapäätöksiksi.
- **Eettinen ja vastuullinen datan käyttö:** ymmärrys datan mahdollisuuksista, rajoituksista ja eettisistä vaikutuksista, erityisesti tekoälyn soveltamisessa.
- **Luova ja kriittinen ajattelu:** kyky yhdistää tietoa eri lähteistä ja muodostaa uusia, innovatiivisia näkemyksiä.
- **Monialaisten tiimien johtaminen ja yhteistyökyky:** datan hyödyntäminen vaatii moniammatillista yhteistyötä, jossa yhdistyvät teknologia, liiketoiminta ja etiikka.
- **Innovatiivisuus ja sopeutumiskyky:** kyky hyödyntää dataa luovasti ja mukautua nopeasti muuttuviin olosuhteisiin auttaa yrityksiä pysymään kilpailukykyisinä.

## Uuden ajan data-ammattilaiset

Teknologian ja tekoälyn roolin kasvaessa organisaatioiden on varmistettava, että niiden datakosysteemit ovat vastuullisesti rakennettuja ja eettisesti käytettyjä. Tämä vaikuttaa myös datan kanssa työskentelevien asiantuntijoiden rooleihin ja osaamisvaatimuksiin.

Perinteisesti data-analyttikot ovat olleet ensisijaisesti loogisesti ajattelevia, syväosaamiseen keskittyviä asiantuntijoita. Tulevaisuudessa heidän on kuitenkin osattava suhteuttaa yksityiskohtainen tieto laajempiin kokonaisuuksiin. Tämä edellyttää uudenlaisten, moniulotteisia näkökulmia yhdistävien datatiimien rakentamista.

Onko tulevaisuuden datatiimi myös organisaation eettinen kompassi? Tätä kysymystä on syytä pohtia, sillä datan mahdollisuuksia ei tule tarkastella pelkästään tehokkuuden näkökulmasta, vaan myös kestäväen kehityksen ja pitkäaikaisten vaikutusten kannalta.

Organisaatiot, jotka ymmärtävät datan arvon ja panostavat sen hallintaan sekä ihmiskeskeiseen teknologian käyttöön, ovat tulevaisuuden liiketoiminnan suunnannäyttäjiä. Tekoälyn rooli on

tärkeä, mutta se ei ole itseisarvo, vaan väline, joka mahdollistaa datan tehokkaan analysoinnin ja hyödyntämisen. Laadukas, eettisesti kerätty ja hyvin mallinnettu data on edellytys kestäväälle ja vastuulliselle liiketoiminnalle.

## Tehtävä 17

Valitse oikea vaihtoehto.

### **Mikä seuraavista kuvaa parhaiten tekoälylukutaidon merkitystä tulevaisuuden liiketoiminnassa?**

- a) Tekoälylukutaito tarkoittaa vain teknistä osaamista ja ohjelmointitaitoja.
- b) Tekoälylukutaito auttaa ymmärtämään tekoälyn toimintaa, hyödyntämismahdollisuuksia ja eettisiä rajoituksia.
- c) Tekoälyn hyödyntäminen liiketoiminnassa ei vaadi lainkaan tekoälylukutaitoa.
- d) Tekoälylukutaito on tärkeää vain datatieteilijöille, mutta sillä ei ole merkitystä muille ammattilaisille.

*Oikea vastaus: b) Datalukutaito on keskeinen taito, koska se auttaa yrityksiä hyödyntämään dataa strategisesti, tekemään parempia päätöksiä ja välttämään harhaanjohtavia johtopäätöksiä. Yritykset, joilla on vahva datalukutaito, voivat käyttää dataa kilpailuetuna.*

### **Miten tulevaisuuden datatiimit eroavat perinteisistä data-analytikkotiimeistä?**

- a) Tulevaisuuden datatiimeillä on vahvempi rooli organisaation eettisenä kompassina.
- b) Perinteiset data-analytikit keskittyivät lähinnä teknisiin yksityiskohtiin, mutta tulevaisuuden datatiimit suhteuttavat tietoa laajempiin kokonaisuuksiin.
- c) Tulevaisuuden datatiimit työskentelevät monialaisissa tiimeissä, joissa yhdistyvät teknologia, liiketoiminta ja etiikka.
- d) Kaikki yllä olevat.

*Oikea vastaus: d) Tulevaisuuden datatiimit eivät ainoastaan analysoi dataa, vaan huomioivat myös eettiset ja liiketoiminnalliset näkökulmat sekä työskentelevät monialaisesti.*

## Data ei ole vain numeroita – se on päätöksiä ja vaikutuksia

Olet nyt käynyt läpi koko kurssin ja saanut kattavan käsityksen siitä, miten data ja tekoäly muokkaavat liiketoimintaa, päätöksentekoa ja yhteiskuntaa. Matka ei kuitenkaan pääty tähän – datan hyödyntäminen ei ole vain tekninen taito, vaan tapa hahmottaa maailmaa ja tehdä parempia valintoja.

Datan arvo ei synny itsestään, vaan vasta silloin, kun sitä käytetään. On yksi asia tietää, että data voi paljastaa trendejä, tunnistaa poikkeamia ja auttaa ennustamaan tulevaa, mutta aivan toinen asia on hyödyntää sitä aktiivisesti omassa työssä ja päätöksenteossa. On helppo ajatella, että data kuuluu vain asiantuntijoille ja analyytikoille, mutta todellisuudessa se on jokaisen ulottuvilla. Kyse ei ole niinkään teknisestä osaamisesta, vaan tavasta esittää kysymyksiä ja nähdä data mahdollisuutena.

Yhteiskunnallisella tasolla data ohjaa päätöksiä yhä laajemmin – oli kyse sitten talouspolitiikasta, terveydenhuollon resursoinnista tai ilmastotoimista. Ymmärtämällä, miten dataa kerätään, käytetään ja tulkitaan, voit vaikuttaa siihen, miten tiedolla johdettu maailma muotoutuu. Datan eettinen käyttö ja läpinäkyvyys ovat keskeisiä periaatteita, jotka määrittävät, kuinka vastuullisesti organisaatiot ja instituutiot toimivat. Kuten olemme oppineet, jotta tämä olisi mahdollista, on äärimmäisen tärkeää ymmärtää datan keruun, säilytyksen ja muokkauksen lähtökohdat.

Tämä kurssi on antanut sinulle välineet käyttää dataa tietoisemmin ja tehokkaammin. Nyt on sinun vuorosi päättää, miten viet opitut asiat käytäntöön. Ehkä se tarkoittaa parempia kysymyksiä raporttien äärellä, rohkeutta kyseenalaistaa oletuksia tai ensimmäistä kokeilua datan visualisoinnissa. Ehkä se merkitsee laajempaa ymmärrystä datan roolista yhteiskunnallisessa päätöksenteossa sekä kriittisempää otetta uutisiin, tilastoihin ja algoritmien vaikutuksiin arjessa.

Datan hyödyntäminen ei ole kertaluontoinen tehtävä, vaan jatkuva ajattelutapa ja toimintamalli. Jokainen analyysi, kysymys ja löydös vie sinua eteenpäin. Datan aikakausi ei ole vain tehokkuuden vallankumous – se on mahdollisuus rakentaa parempia päätöksiä, vastuullisempaa liiketoimintaa ja oikeudenmukaisempaa yhteiskuntaa.

Kurssin keskeinen oppi on selkeä: vain tietoa hyödyntämällä voidaan rakentaa yrityksiä, jotka ovat yhtä aikaa kilpailukykyisiä, eettisiä ja kestäviä. Tulevaisuus kuuluu niille, jotka osaavat yhdistää datan, tekoälyn ja ihmisten osaamisen kokonaisuudeksi, joka ei ainoastaan maksimoi tuottoja, vaan myös tuottaa arvoa yhteiskunnalle ja ympäristölle.

Kiitos, että olet ollut mukana tällä matkalla. Nyt on sinun aikasi ottaa seuraava askel ja hyödyntää dataa omassa työssäsi ja ympäristössäsi – tietoisesti, vastuullisesti ja vaikuttavasti.