

Data-analytiikkaan ja visualisointiin liittyvät perusteet

(Luento 2; 16.10.2024)

Professori, Dosentti

Riitta Hekkala,

riitta.hekkala@aalto.fi

Kauppakorkeakoulu, Aalto yliopisto



Euroopan unionin rahoittama –
NextGenerationEU



Rahoittaja
Jatkuvan oppimisen ja
työllisyyden palvelukeskus

Luennon sisältö

- 1) Datan kerääminen, oikeellisuus, laatu ja hallinta
- 2) Tarinoiden kertominen datan avulla
- 3) Datan analytiikkaan ja visualisointiin liittyvät perusteet ja työkalut

Oppimistavoitteet

- Datan keräämiseen, oikeellisuuteen, laatuun ja hallintaan liittyvät perusteet
- Oppia data-analytiikan ja visualisoinnin liittyvät perusteita ja mahdollisia työkaluja
- Ymmärtää, miten monimuotoisia tapoja on datan tutkimiseen visuaalisesti ja miten tarinoita voidaan kertoa datan avulla

OSA 1: Datan kerääminen, oikeellisuus, laatu ja hallinta



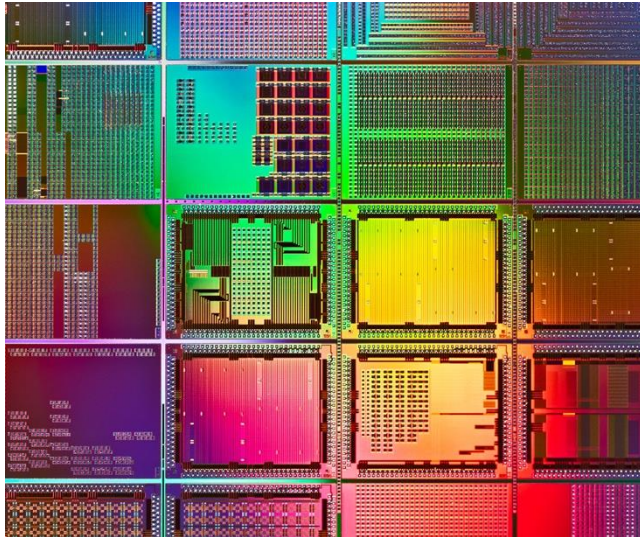
Datan kerääminen ja valmistaminen (Sivula et al. 2023)

- Datan keräämiselle pitäisi aina olla riittävän selkeät perusteet liiketoiminnan kannalta, lakisääteiset asiat huomioiden (esim. GDPR jne.).
- Datan laadusta puhuttaessa huomioitavia asioita ovat
 - Datan tarkkuus, paikkansapitävyys, saatavuus ja datan analysoitavuus.
 - Kenellä on dataan käyttöoikeudet, minkälaisissa rooleissa nämä henkilöt ovat dataan nähden.
 - Datan arvoa lisää se, miten moni yrityksessä osaa/ pystyy käyttämään ja hyödyntämään dataa.

Datan kerääminen ja valmistaminen (Sivula et al. 2023)

- Yritysten tulisi aina ensin panostaa datan keräämiseen ja käsittelyyn liittyvät **prosessit** huolella.
- Datan kerääminen ja varastointi tulisi olla yhtä suunniteltu vaihe kuin yritysten tuotteiden valmistaminen.
- Data ei kehity itsessään eri teknologioita tukevaksi ts. meillä on yhä ollut paljon järjestelmiä, joissa tarvitaan paljon manuaalista työtä, jotta saadaan erilaisia raportteja aikaiseksi.
- Datan säilytys: tietovarastoissa sekä yritysten sisäisissä tietojärjestelmissä, tiedostoissa, tietokannoissa, jotka sijaitsevat joko omissa koneissa tai esim. pilvipalveluissa.

Datan kerääminen ja valmistaminen: preparointi



- **Datan preparointi:** tarkoitetaan raakadatan jalostamista tietovaraksi (datan ryhmittely, yhdistely, standardointi, mahd. puhdistaminen yhdeksi tietojoukoksi).
- Tarkoituksena on tehdä/ saada datasta kelpollista, arvokasta ja käytettävää analytiikkaa varten.
- Datan keräämiseen ja valmistamiseen on sanottu kuluvan suurin osa tiedonanalysoijan työstä (n. 70-80%).

Tiedon oikeellisuus ja laatu

Tiedon tulee olla oikeaa, laadukasta ja helpokäyttöistä, jotta analyysituloksista saadaan tarvittava hyöty.

Tiedon laatua tulee tarkkailla koko sen elinkaaren ajan.

Tiedon eheys ja laadunvarmistus pyritään varmistamaan osana datan puhdistusta (vrt. *ETL-prosessi).

Tietokanta mahdollistaa tietojen tarkkailun jatkuvasti vaihtuvista näkökulmista, jolloin parhaiten löydetään ryhmittelyn ja luokittelun virheitä, epätarkkuuksia tai puutteita.

Tiedon oikeellisuuden ja laadun valvonta ja korjaukset tulisi tehdä tiedontallennuksen tapahtuman käsittelyjärjestelmissä (nk. operatiiviset järjestelmät).

Tiedon oikeellisuus ja laatu

- Operatiivisissa järjestelmissä tiedon oikeellisuus varmistetaan mm.
 - luomalla tiedon kirjaamiseen käytänteitä ja malleja
 - muuntamalla käyttöliittymä tukemaan tietojen syötön oikeellisuutta erilaisten pakollisten kenttien, oletusarvojen avulla.
 - kehittämällä toimintatapoja tiedon ollessa ristiriitaista tai heikkolaatuista varmistetaan tiedon laadullisuus. Tiedon laadun ollessa kunnossa voi itse analyysi alkaa.
 - Datan mahdollinen suuri määrä ja sen jatkuvan lisääntyminen tarkoittaa, että datan puhdistaminen vaatii paljon työtä.
 - *Organisaatiot eivät aina puhdistista koko datamassaa, vaan hyödyntävät vain liiketoiminnan kannalta olennaisen ydintiedon (Master Data). (Valli & Ahlgren 2013)*

Master Data (MD) (esim. Hovi 2015)

- Master Datalla tarkoitetaan pitkäikäistä ja hitaasti muuttuvaa dataa, johon tulee harvoin muutoksia. Usein organisaatioiden keskeinen tieto
 - asiakas- tuote-, henkilöstö- tai organisaatitiedot (esim. asiakasrekisterit).
 - referenssitiedot, kuten erilaiset koodistot (koodilyhenteet, -tunnukset).
- MD-tiedon tallennus ei vaadi jonkin muun tiedon tallennusta ensin ja MD-tietoja ei välttämättä tallenneta päivittäin.



Datan, tiedon ja informaation hallinta



- Datan, tiedon ja informaation hallinnan toimintaketju tulee suunnitella huolella
- Datan luotettavuus ja oikeellisuus: data tallennetaan operatiivisten järjestelmien tietokantoihin, jotka on yhdistetty tietovarastoon (Data Warehouse (DW)).
- Tietovarastoon kerättävä data muunnellaan helpommin käsiteltävään yhteismitalliseen muotoon. Tietovarastoja hyödyntävät taas analyysi- ja raportointijärjestelmät.
- Tietovarastojen raporttien ja analyysien oikeellisuuteen vaikuttaa se, miten ja missä muodossa data on kerätty sekä miten järjestelmän ja tietovaraston välinen integraatio on hoidettu. (Laihonen ym. 2013.)

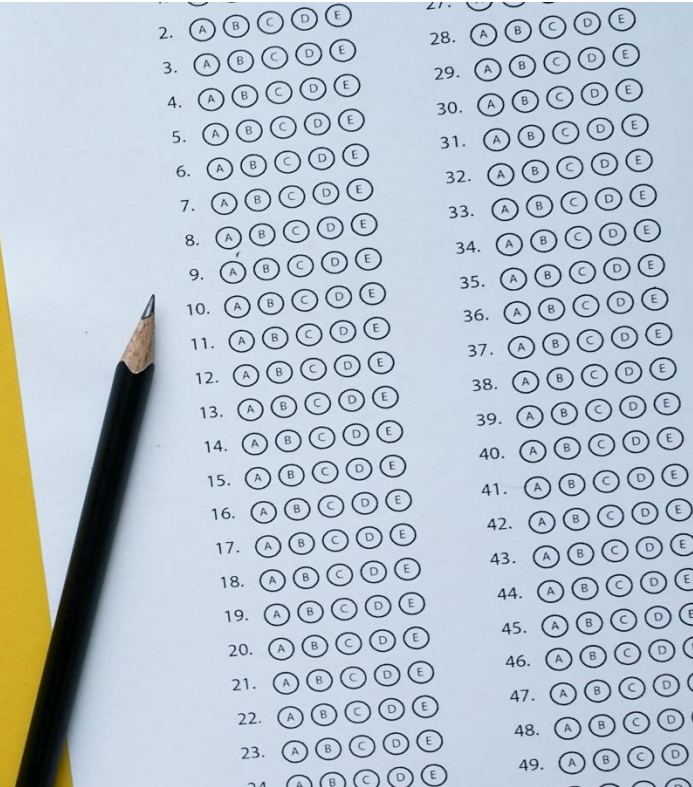
Tiedonhallinnan suunnittelu

- Alkaa organisaation tai sen verkoston ymmärtämisestä
 - Toimiala, sidosryhmät, omistajarakenne, kilpailutilanne toiminta ja sen tavoitteiden (strategiat, visiot).
 - **Viralliset** (organisaatiokartta, koko, kieli, maantieteellinen hajautuneisuus) sekä **epäviralliset** (toimintakulttuuri) rakenteet.
- Lait, standardit ja toimintakohtaiset suositukset.
- Tiedonhallinnan kehittämisen haasteita:
 - Moniulotteisuus ja tiukka sidonnaisuus organisaation ydinprosesseihin.
 - Samoin prosessien ja rakenteiden muuttuessa toimintakin muuttuu, jolla on myös vaikutusta tiedon hallintaan. (esim. Kaario & Peltola 2008, 136.)



Tiedonhallinnan suunnittelu (Datatiede 2018)

- Tiedonhallinta voidaan luokitella tarpeen ja säilytyksen perusteella kolmeen perustasoon.
 - **Arkistoivaa tietoa** tarvitaan lähinnä raportointiin ja sitä tallennetaan määräysten vuoksi.
 - **Ennusteiden ja optimointien avulla erottuvaa tietoa** hyödynnetään edistämään ja parantamaan nykyistä liiketoimintaa.
 - **Innovoivaa tietoa** hyödynnetään löytämään mallinnusten ja visualisointien avulla uusia asioita ja liiketoiminnan mahdollisuuksia.



Data vs. tietokannat

- **Data Warehouse**, DW on raportointia, analytiikkaa ja datan muuta käyttöä varten luotu erillinen tietokanta, nk. analytiikka-alusta (Data Platform), johon hajallaan oleva data yhdistetään helposti saatavilla olevaan muotoon. Datan integroinnin suunnittelussa voidaan hyödyntää erilaisia malleja.
- **Enterprise Data Warehouse**, EDW on organisaatiotason keskitetty tietovarasto, johon voidaan yhdistellä ja integroida organisaation eri tietojärjestelmistä tietoa.
- Pienemmät paikalliset tietovarastot (nk. **datamartit**), johon tieto tuodaan aihekohtaisesti (esim. taloushallinnon tieto) tai osasto- ja organisaatiokohtaisesti henkilöstöosaston tieto. (Hovi et al. 2009)

Yrityksen data vs. liiketoiminta: esim. relaatiotietokannat

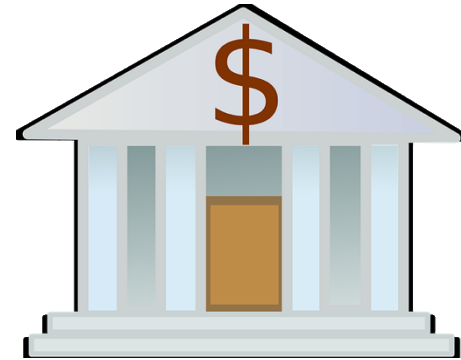
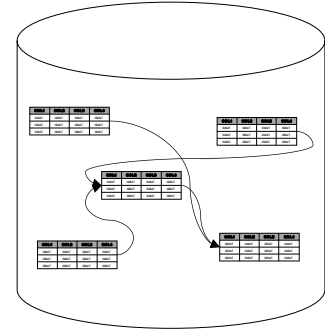
- Joskus yrityksellä voi olla niin paljon erilaista dataa liiketoiminnan kannalta, että tieto joudutaan jakamaan osiin, ts. esim. kaikkea tietoa ei välttämättä voida säilöä yhdessä paikassa.
- Tällöin käytännön tasolla tai teknisellä tasolla, yritykset käyttävät esimerkiksi nk. relaatiotietokantoja.
- Relaatiotietokanta on siis tietokanta, joka muodostaa suhteita eri taulujen välillä, jotka varastoivat dataa erityisiin **tauluihin sisältäen erityiset entiteetit, attribuutit, vierasavaimet jne.**
- Relaatiotietokantoja on ollut 1970-luvulta lähtien ja on toisaalta yksi teknologioista, jotka ihmiset näyttävät ottavan itsestäänselvyytenä näinä päivinä.

Tietokannat - perusteita

- Tietokantaympäristöä varten meidän täytyy ymmärtää **datan yhteydet, datan tyyppi**, jota tietokannassa pidetään, **kuinka dataa käytetään**, ja minkälaisia asioita organisaation tasolla täytyy tehdä tai muuttaa, jotta voimme **hallita dataa** yrityksen näkökulmasta.
- Tärkeää ymmärtää, kuinka organisaatio voi ja pystyy jakamaan datansa tai osan siitä omien liikekumppaneiden ja/ tai kuluttajien kanssa.
- Tietokannat ovat siis hyvin olennainen elementti melkein minkä tahansa järjestelmän ja/ tai sovelluksen ja yrityksen/ organisaation kannalta.

Esimerkkinä: relaatiotietokannat

Let's move forward *with Sisu*



OSA 2: Tarinoiden kertominen datan avulla



Euroopan unionin rahoittama –
NextGenerationEU



Datan merkityksellisuuden ymmärtämisellä pitkä historia



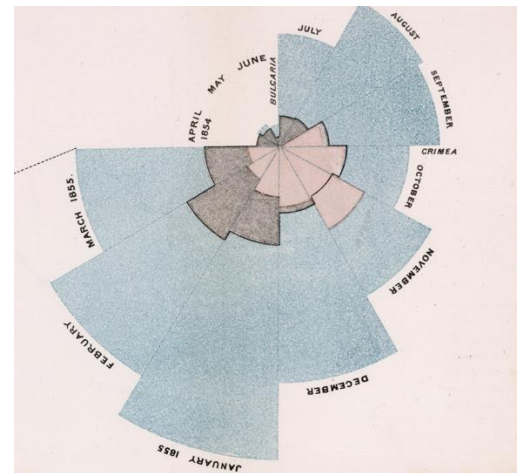
Florence Nightingale
(1820-1910),
Tilastotieteenpioneer

- Data ja sen merkityksellisyys tunnistettu paljon ennen teknologian, eri analyysitekniikoiden ja menetelmien kehittymistä
- ts. datan merkityksellisyys, siitä hyötyminen ei ole uusi asia, mutta kehittynyt teknologia ja eri välineet mahdollistavat meille kuitenkin paljon enemmän
- Florence Nightingale: pidetään tilastotieteen pioneerina

<https://upload.wikimedia.org/wikipedia/commons/1/17/Nightingale-mortality.jpg>

Datan hyödyntäminen – esimerkki jo 1850-luvulta

- Tausta: Krimin sota 1854- vs. brittisotilaiden ”huono hoito sairaaloissa”.
- Nightingale saapui organisoimaan sairaalan käytäntöjä ja kalustoja 38 hoitajan kanssa marraskuussa 1854 Scutariin sotaministerin pyynnöstä.
- Huomioi, että sotilas kuolikin seitsemän kertaa todennäköisemmin sairaalassa (esim. koleraan ja pilkkukuumeseen) kuin taistelussa.
- Nightingale kokosi Turkin sairaaloista ja niiden kuolleisuudesta merkittävää tilastotietoa. Hän havainnollisti tietojaan ympyrädiagrammilla, jota hän nimitti polaaridiagrammiksi.
- Punainen osa edustaa haavoittuneina kuolleita, musta osa muihin syihin kuolleita sininen osa tauteihin kuolleita. Punainen kiila saattaa peittää mustan kiilan (esim. marraskuussa 1854).
- Hänen tilastojensa avulla pystyttiin osoittamaan, että sairaaloiden kuolleisuus oli suurimmillaan tammikuussa 1855, jolloin potilaista kuoli 60 prosenttia jne.



Huhtikuu 1854 –
Maaliskuu 1855

<https://upload.wikimedia.org/wikipedia/commons/1/17/Nightingale-mortality.jpg>

Data-analytiikka ja tarinat

- Tärkeää ajatella tarinaa, ennen kuin luo data visualisoinnin. **Data ei itsessään kerro vielä mitään, eikä myöskään visualisoinnit**
- *“Analysts who care whether their work is implemented— whether it changes decisions and influences actions—care a lot about this issue and devote a lot of time and effort to it. Analysts who don’t care about such things believe that the results ‘speak for themselves,’ and don’t worry much about communications.” (Davenport, 2013, p. 72)*



Data-analyysin prosessi (Valli & Ahlgren 2013)

- 1) Suunnittelu eli ongelman määrittäminen kysymysten kautta
- 2) Valmistelu eli datan kerääminen ja valinta
- 3) Esikäsittely eli puhdistus, muokkaus, suodatus, täydentäminen, standardointi
- 4) Datan analysointi tutkimalla eli luokittelu, ryhmittely, mallinnus sekä korrelaatiot
- 5) Data-analyysien toteutus eli visualisointi
- 6) Tulosten tulkinta eli raporttien vertaaminen menneeseen ja ennustaminen
- 7) Data-analyysien hyödyntäminen.

Tarinoiden kertominen datan avulla

- Lukuisia artikkeleita siihen, miten datan avulla kertoa tarina:
 - 1. Etsi vakuuttava narratiivi.** Tosiasioiden esittämisen ohella luo yhteydet niiden välillä.
 - 2. Mieti yleisöä/ kohderyhmää.** Mitä he tietävät aiheesta?
 - 3. Ole objektiivinen.** Visualisaation pitäisi aina olla objektiivinen ja tietyissä tilanteissa voi olla, että on esteellinen (biased) vaikuttamaan. Pitäisi aina pohjautua tosiasioihin, ei siihen mitä halutaan sanoa
 - 4. Älä ole valikoiva aineiston suhteen,** ts. Mitä sisällytät tai et sisällytä, ennen kuin olet tarpeeksi luottavainen, että annat parhaan esitysmuodon siitä, mitä aineisto “sanoo”.
 5. Lopuksi, **editoi** – huolellisesti arvioi, että data tulee esitellyksi

OSA 3: Data-analytiikkaan ja visualisointiin liittyvät perusteet ja työkalut



Mitä data-analytiikalla tavoitellaan?

- **Tavoite:** raakadatan erittelemistä ja jäsentelemistä erilaisiin käyttötarkoituksiin.
- **Yksityiskohtaisempia tavoitteita...**
 - Yhtenäinen tietoperusta, ajantasainen tieto.
 - Päätöksenteon hajauttaminen ja nopeuttaminen organisaatioissa.
 - Mahdollistaa uusien toimintatapojen, palveluiden ja uusien liiketoimintaideoiden kehityksen.
 - Auttaa ymmärtämään asiakasta ja kehittämään parempaa asiakaskokemusta.



Miksi data-analytiikkaa kannattaa tehdä?

- Data-analytiikan avulla dataa voidaan hyödyntää päätöksenteon tukena, ja selvittää pitääkö mahdolliset omat olettamukset ja näkökulmat liiketoiminnasta paikkansa.
- Data-analytiikka auttaa näkemään, että mitkä toimenpiteet mahdollisesti johtavat parempiin lopputuloksiin.
- Datasta voi tehdä eri lähestymistavoin myös uusia löydöksiä, joita ei omien päätelmien kautta ole tullut/ tule esille tai ajatelleeksi.



Data-analyysi – ”Jos et voi mitata, et voi hallita” (Drucker)

- Data-analyysilla tarkoitetaan menetelmiä, joiden avulla datasta muokataan informaatiota päätöksenteon tueksi. Tähän liittyy myös tiedon esittäminen visuaalisesti sekä eri mallinnusten, todennäköisyyksien tai ennusteiden muodostaminen.
- Hyvä mittari onnistumiselle on vertaileva, ymmärrettävä ja se mittaa mm. asioiden syy-seuraussuhdetta. Lisäksi mitatun muutoksen perusteella, voidaan tehdä päätelmiä, että mitä tehtäisiin eri tavalla. (Croll & Yoskovitz 2019.)

Data-analyysin arvo

- Datan arvo syntyy sitä hyödynnettäessä. Data-analytiikan yksi keskeinen tavoite aina on, että se auttaa tekemään liiketoiminnallisesti parempia päätöksiä.
- Olennaisinta data-analytiikassa on **ymmärrys ja käsitys**, siitä
 - Mitä dataa saa kerätä ja kuinka sitä muunnetaan?
 - Hyödynnettävistä menetelmistä ja tekniikoista
 - Tuloksien hyödynnettävyydestä esim. tavoitteiden määrittelyssä, strategiassa, prosessin seurannassa, päätöksenteossa ja tulosten jakamisessa
 - Tiedon käytön mahdollisuuksista.

Data-analytiikka prosessina



- Vaihtoehtoisia prosesseja, työkaluja ja menetelmiä
 - Data-analytiikan tuloksista saadaan enemmän ennakoitavia, ymmärrettäviä
 - Helpommin hyödynnettäviä ja uudelleen käytettäviä.
- Data-analytiikka voi tehdä myös ad-hoc tyyliin, ts. selvittämällä mitä kulloinkin haluaa datasta erityisesti tietää ilman tarkempaa prosessia.
- Muodolliset prosessikuvaukset voivat auttaa kuitenkin ymmärtämään, mitä kaikkea data-analytiikkaan liittyy ja mitä analysoinnissa olisi hyvä ottaa huomioon.

Data-analytiikan perustasot

Kuvaileva analytiikka

Ennakoiva analytiikka

Ohjaileva analytiikka

Diagnosoiva analytiikka



Data-analytiikan eri muodot (Davenport & Harris, 2017)

- **1. Deskriptiivinen analytiikka:** “Mitä liiketoiminnassa on tapahtunut? Kuinka paljon, kuinka usein? Mitä toimenpiteitä tarvitaan?”
- **2. Diagnostiivinen analytiikka:** “Miksi näin on tapahtunut?”
- **3. Prediktiivinen analytiikka:** “Mitä tapahtuu jos nämä trendit jatkuvat? Mitä tapahtuu seuraavaksi? Mitä todennäköisesti tapahtuu tulevaisuudessa?”
- **4. Preskriptiivinen analytiikka:** “Mitä tapahtuu, jos kokeillaan tätä? Mikä on parasta mitä voi tapahtua? Mitä toimenpiteitä tulisi tehdä?”
- **5. Ennaltaehkäisevä analytiikka:** “Kuinka voimme välttää tilanteen? Mitä tulisi tehdä enemmän?”
- **6. Autonominen analytiikka:** “Mitä voimme oppia datasta?”

1. Deskriptiivinen (kuvaileva) analytiikka (Datatiede 2018; Jussila et al. 2019)

- Selvitetään, että mitä on tapahtunut hyödyntäen tilastollisia menetelmiä, kuten summaa, keskiarvoa ja prosenttimääräisiä muutoksia.
- Visuaalisessa raportissa esitetään esim. organisaation tuoton, rahoituksen, tuotannon tai myynnin muutoksia.
- Kuvailevan analytiikan menetelmiä ovat datan kategorisointi ja luokittelu.

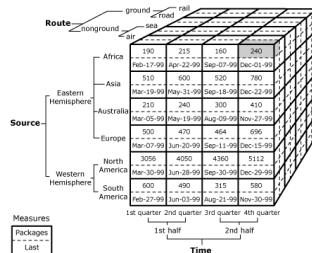
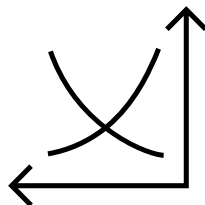
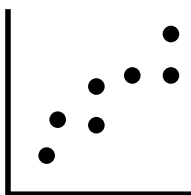


2. Diagnosoiva analytiikka (Jussila et al. 2019)

- Tapahtumien syy-yhteydet, ts. miksi jotain on tapahtunut.
- Asioiden keskinäiset vaikutukset etsitään yhdistelemällä eri tietolähteitä ja tilastoja.
- Tuloksena saadaan tapahtumien eri yhteyksien vaikutusten todennäköisyyksien arvo.



- Diagnosoivan analytiikan menetelmiä ovat mm. korrelaatiot, Online Analytical Processing (OLAP-kuutiot) sekä tiedonlouhinta. (Datatiede 2018.)



3. Ennakoiva (prediktiivinen) analytiikka: tapahtuu historian pohjalta (esim. Jussila et al. 2019)

- Datasta pyritään löytämään erilaisia toistuvia malleja ja datajoukkojen välisiä suhteita ennustamaan esim. asiakaskäyttäytymistä, ostotrendejä tai tuotantoketjun vaatimuksia.
- Menetelmiä ovat tilastolliset menetelmät, kuten lineaarinen ja logistinen regressio, faktorianalyysi, klusterointi, joita hyödynnetään trendien ymmärtämiseen ja tulevien tapahtumien ennustamiseen.
- Mallien etsintään käytetään tiedonlouhintaa hyödyntäen neuroverkkoja ja geneettisiä algoritmeja.



4. Preskriptiivinen eli ohjaileva analytiikka (esim. Jussila et al. 2019)

- Selvittää kuinka saadaan aikaiseksi haluttu tulos eli suosittelee mitä tulisi tehdä. Tämän avulla tuotetaan ehdotuksia tulevien mahdollisuuksien hyödyntämiseen tai tulevien riskien vähentämiseen.
- Ohjaava analytiikka käsittelee tietoa jatkuvasti ja automaattisesti verraten ja ennustaen toimenpiteiden vaikutuksia.
- Tämän avulla parannetaan ennustamisen tarkkuutta ja tuotetaan parempia toimenpide-ehdotuksia.
- Vaihtoehtojen suuri määrä: optimoinnin avulla pyritään löytämään paras mahdollinen vaihtoehto.

Muita analytiikan tasoja (Jussila et al. 2019)

- **5. Ennaltaehkäisevä analytiikka**
 - *“Kuinka voimme välttää tilanteen?”*
 - *“Mitä tulisi tehdä enemmän?”*
- **6. Autonominen analytiikka**
 - *“Mitä voimme oppia datasta?”*

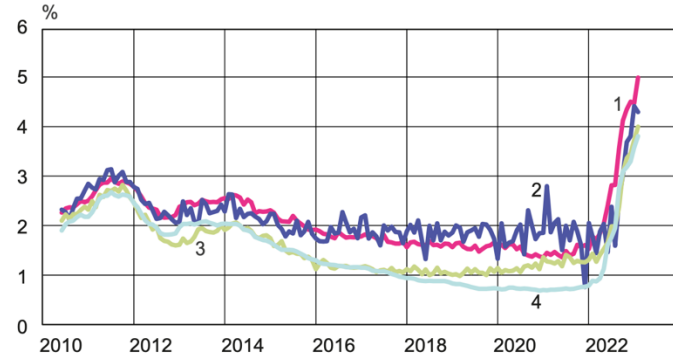


Esimerkkinä: Suomen Pankin ajankohtaisia asioita taloudesta

- Kuvaileva (Mitä liiketoiminnassa on tapahtunut? Kuinka paljon, kuinka usein? Mitä toimenpiteitä tarvitaan?)
- Esim. **Mitä tapahtunut:** korkojen äkillinen nousu paljastaa alhaisten korkojen aikana otettuja riskejä
- **Mahd. syy-yhteydet:** Vauhdittunut inflaatio ja korkojen jyrkkä nousu ovat muuttaneet lyhyessä ajassa lainanhakijoiden, velallisten, pankkien ja talouspolitiikan toimintaympäristöä.
- Markkinakorot ovat nousseet euroalueella vuoden aikana nollassa useisiin prosentteihin.

Uusien asunto-, yritys- ja asuntoyhteisölainojen korot nousseet Suomessa nopeasti

1. Kotitalouksien uudet lainat yhteensä
2. Yritysten uudet lainat
3. Asuntoyhteisöjen uudet lainat
4. Kotitalouksien uudet asuntolainat



Lainat Suomen rahalaitoksilta, pl. tili- ja korttiluotot.
Lähde: Suomen Pankki.

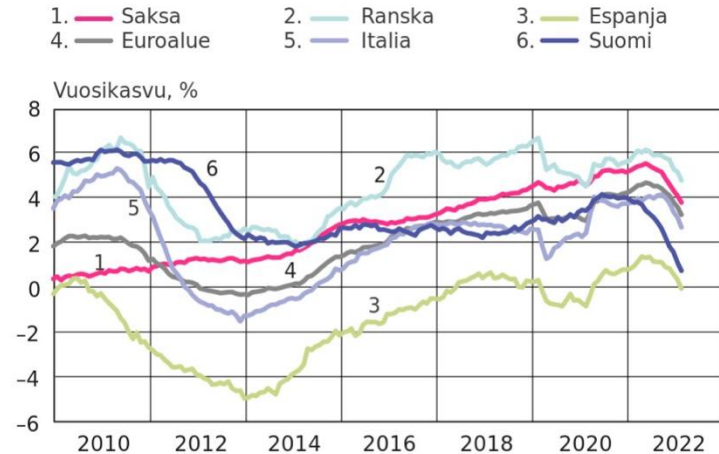
24.4.2023

© Suomen Pankki / (© Bank of Finland)

Esimerkinä: Kotitalouslainojen kasvu euromaissa ja Suomessa

- Korkeiden korkojen nousu on kääntänyt rahoitussuhdanteen laskuun Euroopassa. Kiinteistöjen hintojen pitkään jatkunut nousu on monissa maissa taittumassa, ja yritysten ja kotitalouksien luotonannon kasvuvauhdit hiipuvat.
- Kotitalouslainojen nostot ovat vähentyneet Suomessa erityisen jyrkästi.
- Rahoitussuhdanteen nousuvaiheessa mahdollisesti syntyneet haavoittuvuudet – kuten asuntojen hintojen yliarvostus tai rakentamisen ylitarjonta – voivat johtaa rahoitussuhdanteen laskuvaiheessa rahoitusjärjestelmän vakaviin häiriöihin.

Kotitalouslainojen kasvu hidastunut nopeasti suurissa euromaissa ja etenkin Suomessa



Lähde: Euroopan keskuspankki.

24.4.2023

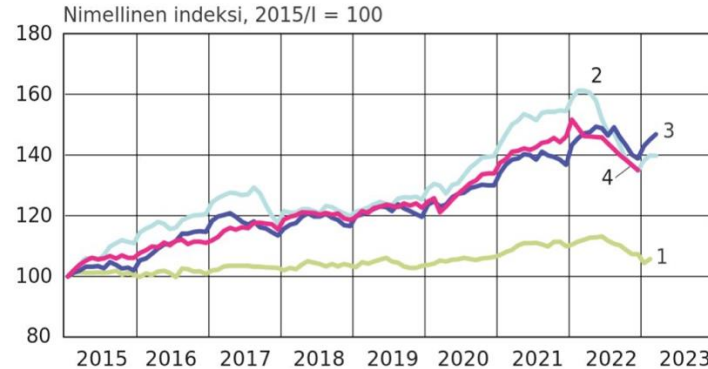
© Suomen Pankki / (© Bank of Finland)

Esimerkkinä: Asuntojen hintojen lasku Pohjoismaissa

- Asuntojen hinnat ovat alkaneet laskea Pohjoismaissa
- Hinnat ovat laskeneet Suomessa vuoden 2022 huipusta (n. 7,5 %).
- Pohjoismaisen vertailun mukaan hintojen lasku on ollut jyrkintä Ruotsissa, jossa asuntojen hinnat nousivat erityisen voimakkaasti pandemian aikana.
- Korkeiden korkojen nousu ja rahoitusolojen kiristyminen vaikuttavat velallisiin, sijoittajiin, rahoitusalan toimijoihin ja niiden kautta rahoitusvakauteen.

Asuntojen hinnat ovat alkaneet laskea kaikissa Pohjoismaissa

1. — Suomi 2. — Ruotsi 3. — Norja 4. — Tanska



Suomi: vanhat osakeasunnot. Ruotsi, Tanska ja Norja: kaikki asunnot. Tanskan aineisto kvartaaleittain vuodesta 2021 lähtien. Lähde: Tilastokeskus, Valueguard, Statistics Denmark, Eiendom Norge, Macrobond ja Suomen Pankin laskelmat.

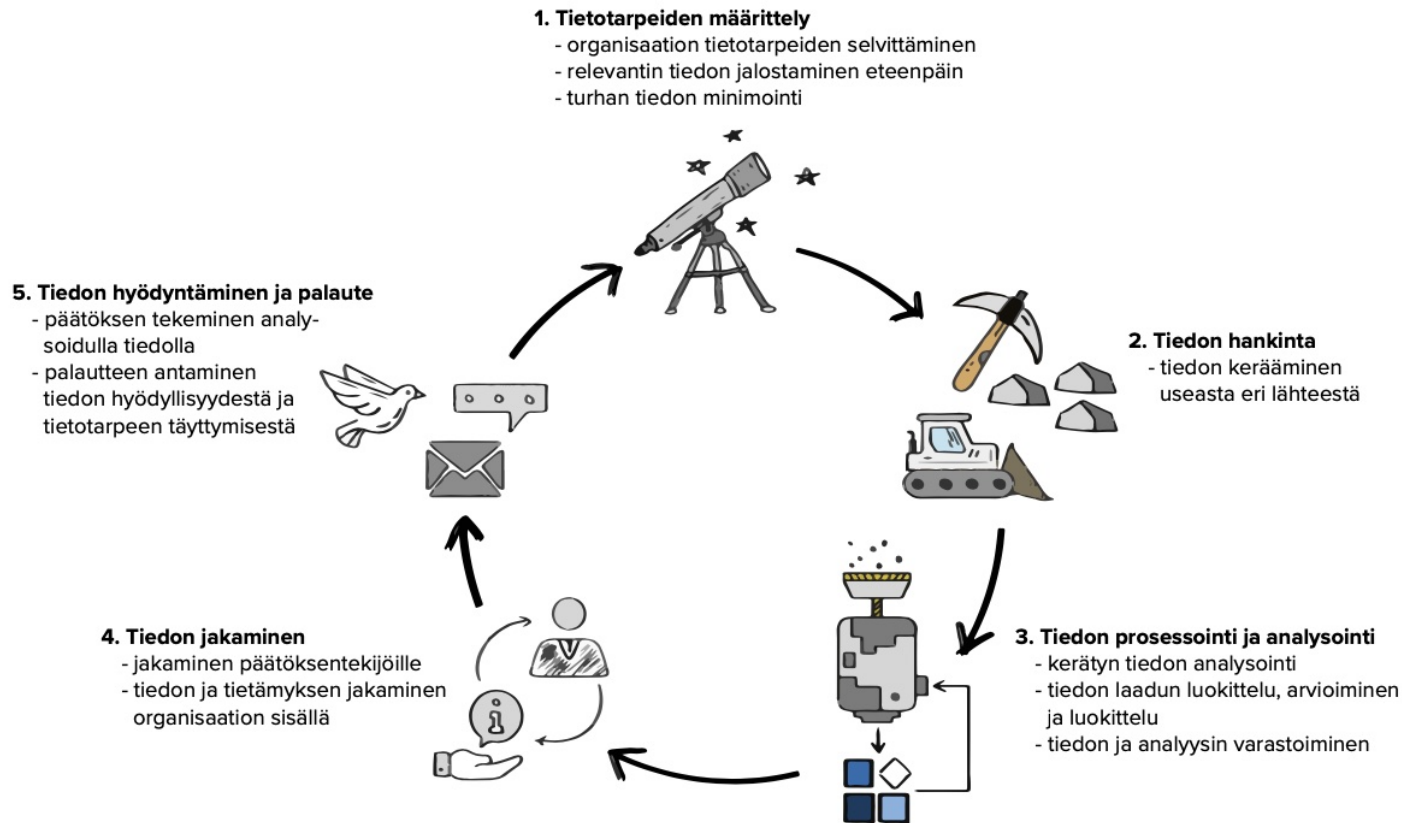
24.4.2023

© Suomen Pankki / (© Bank of Finland)

Liiketoimintatiedon hallinta (Laihonen et al. 2013)

- Liiketoimintatiedon hallinnalla tarkoitetaan toimintaa, jonka avulla organisaatio kerää, analysoi ja hyödyntää oman liiketoiminnan kannalta olennaista dataa.
- Liiketoimintatiedon päämääränä on kerätä tarpeeksi dataa eri lähteistä talteen jatkokäyttöä varten.
- Tarkoituksena on yhdistellä datasta muodostuneita irrallisia tiedonpalasia ja tätä kautta auttamaan päätöksentekijöitä organisaation sisällä.
- Tämän tiedon avulla päätöksentekijät pystyvät ennakoimaan ja tunnistamaan organisaation riskitekijöitä ja muuttamaan liiketoimintaan tuottavampaa ja parempaan suuntaan.

Liiketoimintatiedon hallinnan prosessimalli (Jussila et al. 2019, s. 3)




Datan visualisoinnin perusteita

A? Aalto-yliopisto
Aalto-universitetet
Aalto University



Euroopan unionin rahoittama –
NextGenerationEU

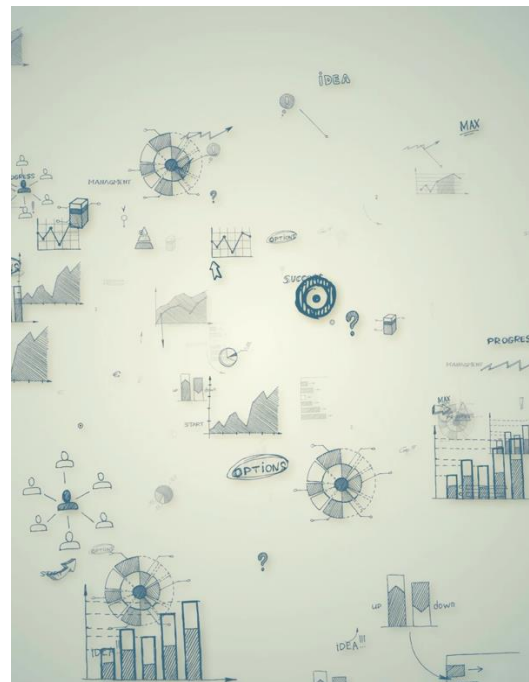
Rahoittaja
 Jatkuvan oppimisen ja
työllisyyden palvelukeskus

Milloin dataa kannattaa visualisoida?

- **Monimutkainen data:** Kun dataa on paljon, sen esittäminen kaavioiden tai diagrammien avulla tekee datasta helpommin ymmärrettävää ja selkeämpää.
- **Suhteiden esittäminen:** Kaavioiden ja diagrammien avulla voi esittää tietojen välisiä suhteita, kuten korrelaatioita ja/ tai ajan suhteita.
- **Tietojen vertailu:** Tietojen vertailu voi olla helpompaa, kun niitä esitetään graafisesti. Tämä voi auttaa päätöksenteossa, ts. tekee sitä helpommaksi ja tehokkaammaksi.

Milloin dataa kannattaa visualisoida?

- **Trendit ja kehitys:** Kaaviot ja diagrammit ovat hyödyllisiä, kun halutaan seurata tietojen kehitystä ajan yli.
- Tiedon visualisointi on hyödyllistä silloin, kun haluat tehdä tietojen esittämisestä selkeämpää, helpommin ymmärrettävää ja muistettavampaa.
- Se on erityisen hyödyllistä suurten ja monimutkaisten tietovarantojen kanssa vrt. kuitenkin alat

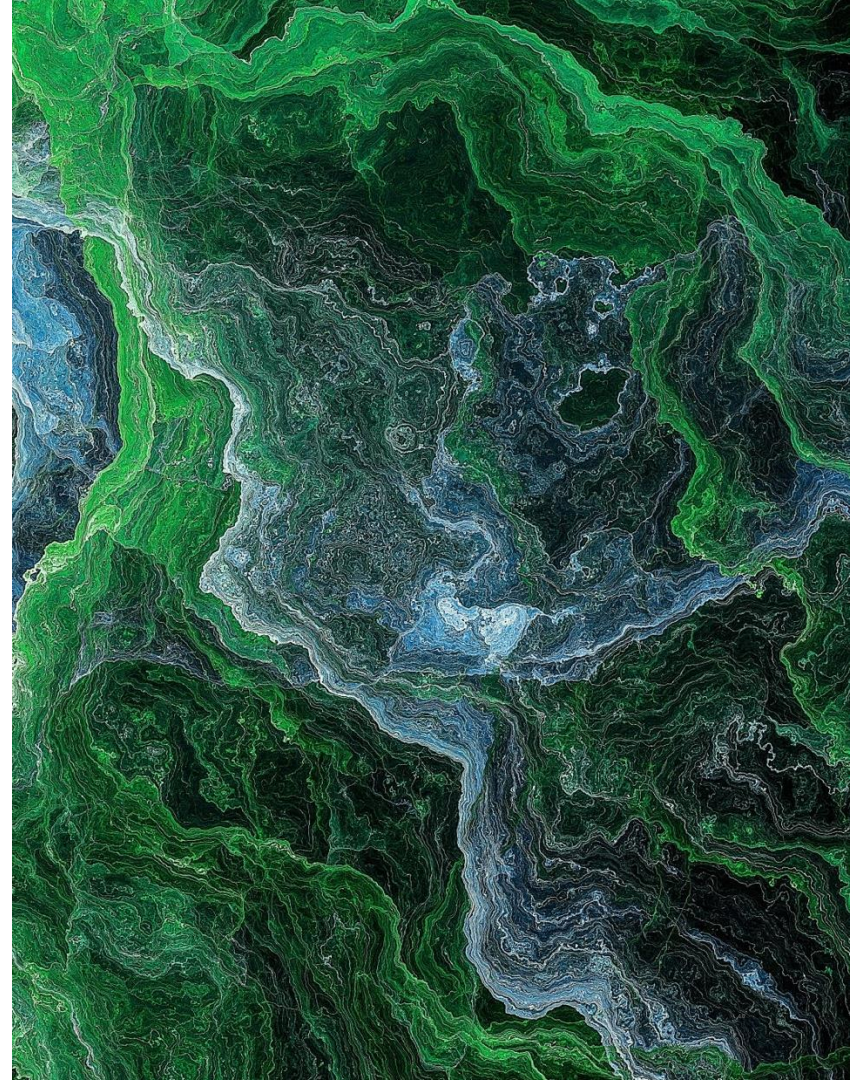


Milloin data kannattaa esittää taulukkona ja milloin kaavioina?

- **Taulukko** sopii hyvin tietojen esittämiseen numeerisessa muodossa, kun tärkeintä on esittää tarkkaa ja yksityiskohtaista tietoa.
 - Esimerkiksi tilastot, budjetit tai laskentatiedot.
- **Kaaviot** sopivat paremmin tietojen esittämiseen visuaalisessa muodossa, ja ne ovat hyviä kuvatessa suhteellisia arvoja ja suuntauksia. Kaaviot ovat erityisen hyviä esittämään suuria tai monimutkaisia tietosarjoja, joita on vaikea ymmärtää taulukoista.
 - Esimerkiksi linjakaaviot, pylväskaaviot tai piirakkakaaviot ovat hyviä vaihtoehtoja.

Kaaviotyypit tiedon visualisointiin

- Pylväskaavio
- Viivakaavio
- Piirakkadiagrammi
- Hajontakuvio
- Histogrammi
- Laatikko–jana-kaavio



Konkreettisia vinkkejä tiedon visualisointiin - kaavioiden esittäminen

- **Vältä tarpeetonta informaatiota:** tarpeettoman informaation sisällyttäminen visualisointiin saattaa vääristää tietojen välittämää merkitystä. Se saattaa myös johtaa tarpeettomiin yksityiskohtiin keskittymiseen.
- **Käytä selkeitä fontteja:** Tiedon visualisoinnissa kannattaa käyttää selkeitä fontteja, sillä ne helpottavat tietojen lukemista ja ymmärtämistä. Liian pieni tai kalligrafisesti muotoiltu fontti voi tehdä tekstistä hankalaa lukea ja häiritä visuaalista esitystä.

Konkreettisia vinkkejä tiedon/datan visualisointiin

- **Lyhyt ja ytimekäs teksti** auttaa keskittymään tärkeimpään tietoon. Lisäksi lyhyt teksti auttaa pitämään kaavion yksinkertaisena ja selkeänä.
- **Symbolit** ovat tärkeitä tiedon visualisoinnissa, sillä ne helpottavat ymmärrystä ja edistävät tiedon muistamista. Tärkeää on valita symbolit, jotka ovat selkeitä, estävät väärinkäsityksiä ja ymmärrettäviä kaikille, jotka katsovat visualisointia.
- **Värejä** kannattaa käyttää tiedon visualisoinnissa tarkoituksellisesti ja harkiten, jotta ne todella tukevat tietojen esittämistä ja tekevät visualisoinnista helposti ymmärrettävän. Väreillä on merkittävä vaikutus sen selkeyteen ja tehokkuuteen.

Konkreettisia vinkkejä tiedon/ datan visualisointiin



Järjestä tiedot tärkeysjärjestykseen.



Vältä tarpeetonta yksityiskohtaisuutta.



Käytä animaatioita harkiten



Miten voin varmistaa, että visualisoitava tieto on ymmärrettävää

Millaisia työkaluja tiedon visualisointiin voi käyttää?

- **Tableau** on johtava tiedon visualisointityökalu, joka on suunniteltu erityisesti liiketoiminnan tarpeisiin. Se tarjoaa laajan valikoiman visuaalisia esitystapoja ja mahdollistaa tiedon yhdistämisen ja analysoinnin. (2003->)
- **PowerBI** on Microsoftin tarjoama tiedon visualisointityökalu, joka on suunniteltu helpottamaan tiedon esittämistä ja analysointia. Se tarjoaa erilaisia visuaalisia esitystapoja ja mahdollistaa tiedon yhdistämisen ja analysoinnin. (2015->)
- **QlikView** on tiedon visualisointityökalu, joka on suunniteltu tarjoamaan tehokkaan tiedonkäsittelyratkaisun. Se tarjoaa monipuolisia visuaalisia esitystapoja ja mahdollistaa tiedon yhdistämisen ja analysoinnin. (1994->)

Millaisia työkaluja tiedon visualisointiin voi käyttää?

- **D3.js** on avoimen lähdekoodin JavaScript-kirjasto, joka mahdollistaa ja tarjoaa laajan valikoiman visuaalisia esitystapoja ja mahdollistaa tiedon yhdistämisen ja analysoinnin. (2011->)
- **Matplotlib** on yksi suosituimmista ja monipuolisimmista Python-kirjastoista tiedon visualisointiin. Se tarjoaa monia graafisia esitysmuotoja, kuten histogrammeja, kuvaajia ja diagrammeja (2003->).
- **Microsoft Excel** on perinteinen tiedon visualisointityökalu, joka tarjoaa mahdollisuuden luoda yksinkertaisia kaavioita ja taulukoita. Vaihtoehtoisia taulukko-ohjelmia Microsoft Excelille ovat Google Sheets ja Open Office Calc. (1987->)

Lähteet (artikkelit/ kirjat)

- Datatiede 2018; Davenport, T., & Harris, J. (2017). Competing on analytics: Updated, with a new introduction: The new science of winning. Harvard Business Press.
- Croll, A & Yoskovitz, B (2019). LEAN Analytics, Use Data to Build a Better Startup Faster. O'Reilly.
- Hovi, A. Hervonen, H. Koistinen, H. (2009). Tietovarastot ja business intelligence WSOY, Jyväskylä
- Jussila, J., Saari, J., Närhi, J. & Järvenpää, A.-M. (2019). TULEVA, Data-analytiikan opas pk- yrityksille 2.0. https://www.hamk.fi/wp-content/uploads/2019/01/TULEVA_kevytopas_data-analytiikka.pdf

Lähteet (artikkelit/kirjat)

- Laihonen, H., Hannula, M., Helander, N., Ilvonen, I., Jussila, J., Kukko, M., ... Yliniemi, T. (2013). Tietojohtaminen. Tampere: Tampereen teknillinen yliopisto, Tietojohtamisen tutkimuskeskus Novi.
- Valli, K. & Ahlgren, S. (2013). Informaatiosta kilpailuetua teollisuusyrityksiin.
https://teknologiateollisuus.fi/sites/default/files/file_attachments/elinkeino-politiikka_digitalisaatio_teollinen_internet_informaatiosta_kilpailuetua.pdf

Lähteet (www-sivuja)

- <https://hbr.org/2013/04/how-to-tell-a-story-with-data>
- <https://hbr.org/2013/04/how-to-tell-a-story-with-data>
- Suomen pankin esimerkit:
<https://www.eurojatalous.fi/fi/2023/1/korkojen-nousu-testaa-velallisten-sijoittajien-ja-rahoitusjarjestelman-kestavyytta/>
- + useassa kalvossa erikseen sivu

A field of red poppies is shown at sunset. The sun is low on the horizon, creating a warm, golden glow. The poppies are in various stages of bloom, with some fully open and others as buds. A semi-transparent white box is overlaid on the left side of the image, containing text.

Kiitos!
Kysymyksiä?

riitta.hekkala@aalto.fi