



Vaasan yliopisto  
UNIVERSITY OF VAASA

Niina Kankaanpää

**Business Intelligence ja data-analytiikka  
eteläpohjalaisissa  
yrityksissä**

Laskentatoimen ja rahoituksen  
akateeminen yksikkö  
Laskentatoimen ja tilintarkastuk-  
sen pro gradu -tutkielma  
Laskentatoimen ja tilintarkastuk-  
sen maisteriohjelma

Vaasa 2022

---

**VAASAN YLIOPISTO****Laskentatoimen ja rahoituksen yksikkö**

<b>Tekijä:</b>	Niina Kankaanpää	
<b>Tutkielman nimi:</b>	Business Intelligence ja data-analytiikka eteläpohjalaisissa yrityksissä	
<b>Tutkinto:</b>	Kauppatieteiden maisteri	
<b>Oppiaine:</b>	Laskentatoimen ja tilintarkastuksen maisteriohjelma	
<b>Työn ohjaaja:</b>	Marko Järvenpää	
<b>Valmistumisvuosi:</b>	2022	<b>Sivumäärä:</b> 102

---

**TIIVISTELMÄ**

Saatavilla olevan datan määrä on maailmassa kasvanut räjähdysmäisesti ja siihen pohjautuvaa data-analytiikkaa on alettu pitämään yrityksissä kilpailukykyä tuovana menetelmänä. Tämän tutkielman tavoitteena on tuottaa tiedeyhteisölle ajantasainen kuva siitä, käyttävätkö yritykset Etelä-Pohjanmaalla Business Intelligenceä tai data-analytiikkaa liiketoimintansa päätöksenteon tukena. Tutkielmassa kartoitetaan nykytilannetta ja selvitetään mistä yritykset hakevat dataa ja mitä työkaluja he käyttävät sen analysoinnissa. Varsinaisena tavoitteena on selvittää, millä tasolla yritykset Etelä-Pohjanmaalla kokevat olevansa BI:n ja data-analytiikan suhteen ja näkevätkö he nykyisen tason riittävänä. Tutkielmassa selvitetään myös yritysten tulevaisuuden näkymiä ja kehittämistarpeita data-analytiikan saralla.

Tutkielman teoriaosuuden ensimmäisessä osassa esitellään dataa ja sen roolia päätöksenteon välineenä. Samassa avataan data-analytiikan käsitteen taustoja ja verrataan sitä Business Intelligencen käsitteeseen. Toisessa osassa paneudutaan enemmän data-analytiikkaan ja sen rooliin yrityksissä. Teoriaosuudessa esitellään aikaisempia tutkimuksia aiheiden ympärillä.

Tutkielman empiirinen osuus toteutettiin puolistrukturoiduilla teemahaastatteluilla. Tutkimusjoukkona toimivat Etelä-Pohjanmaata kotipaikkanaan pitävät yritykset, joiden liikevaihto vuositasolla oli yli 100 miljoonaa. Kyseessä oli kvalitatiivinen tutkielma. Haastattelujen kohteeksi valittiin 6 haastateltavaa, jotka toimivat yrityksissä erilaisissa liiketoiminnan johto- ja päällikkötason tehtävissä.

Tutkielman tulokset osoittavat Eteläpohjalaisten yritysten olevan melko tyytyväisiä tämänhetkiseen data-analytiikan tasoon. Suurin osa yrityksistä on analytiikan hyödyntämisessä tasolla Business Intelligence, kun taas muutamassa hyödynnetään jo hieman edistyksellisempää data-analytiikkaa. Tulevaisuuden osalta data-analytiikkaa uskottiin kehitettävän ja sen hyödyntämiselle nähtiin erilaisia kohteita yritysten liiketoiminnassa.

---

**AVAINSANAT:** Big data, Business intelligence, Data-analytiikka

## Sisällys

1	Johdanto	6
1.1	Tutkielman tausta	6
1.2	Tutkielman aihe ja tavoitteet	7
1.3	Tutkielman rajaukset	8
1.4	Tutkielman rakenne	9
2	Analytiikan käsitteet	11
2.1	Analytiikka	11
2.2	Data	12
2.2.1	Big Data	14
2.2.2	Datan laatu ja luotettavuus	18
2.2.3	Data päätöksenteossa	19
2.3	Business intelligence	22
2.4	Data-analytiikka	24
2.4.1	Data-analytiikan käsite	25
2.4.2	Business intelligencen ja data-analytiikan erot	29
2.4.3	Data-analytiikan työkalut	33
3	Data-analytiikka yrityksissä	36
3.1	Liiketoiminnan haasteet ja data-analytiikka	37
3.2	Data-analytiikka päätöksenteossa ja data-analytiikan haasteet	38
3.3	Data-analytiikan tulevaisuuden näkymät	43
3.4	Teoreettinen tarkastelu ja tutkimuskysymykset	44
4	Tutkielman toteutus	46
4.1	Tutkimusmetodologia	46
4.2	Aineiston hankinta	47
4.3	Aineiston analysointi	49
5	Tutkielman tulokset	51
5.1	Taustatiedot ja tutkielman lähtökohdat	51

5.2	Datan lähteet ja analysointityökalut	54
5.3	Datan ja analysoinnin olennaisuus ja riittävyys sekä haasteet	61
5.3.1	Datan ja analysoinnin olennaisuus ja riittävyys	61
5.3.2	Liiketoiminnan ja data-analytiikan haasteet	72
5.4	Data-analytiikan riittävyys ja tulevaisuuden näkymät	77
5.5	Yhteenveto tuloksista	82
6	Johtopäätökset	87
6.1	Yhteenveto tutkielmasta	87
6.2	Tutkielman luotettavuus	92
6.3	Jatkotutkimusehdotukset	93
	Lähteet	95
	Liitteet	101

## Kuviot

<b>Kuvio 1</b>	Strukturoitu, semistrukturoitu ja strukturoimaton data.	13
<b>Kuvio 2.</b>	Big datan 6 V:tä.	16
<b>Kuvio 3.</b>	Data-analytiikan menetelmät.	28
<b>Kuvio 4.</b>	Data-analytiikan elämänkaari.	29
<b>Kuvio 5.</b>	Business intelligenen ja analytiikan erot.	31
<b>Kuvio 6.</b>	Data-analytiikkapohjainen päätöksentekoprosessi.	42

## Taulukot

<b>Taulukko 1.</b>	Datalähteet ja analysointityökalut.	82
<b>Taulukko 2.</b>	Datan ja analysoinnin olennaisuus ja riittävyys.	84
<b>Taulukko 3.</b>	Data-analytiikan laajentaminen ja tulevaisuuden näkymät.	85

## Lyhenteet

AI	Artificial Intelligence (tekoäly)
BA	Business Analytics (liiketoiminta-analytiikka)
BD	Big Data (massadata)
BDA	Big Data Analytics (massadata-analytiikka)
BI	Business Intelligence (liiketoimintatiedon hallinta)
BI&A	Business Intelligence and Analytics (liiketoimintatiedon hallinta ja analytiikka)
CRM	Customer Relationship Management (asiakkuuksien hallintajärjestelmässä)
ERP	Enterprise Resource Planning (toiminnanohjausjärjestelmä)
IoT	Internet of Things (esineiden ja asioiden internet)
IT	Information technology (tietotekniikka)
KPI	Key Performance Indicator (suorituskykymittari)
MD	Master Data (ydintieto)
MDM	Master Data Management (ydintiedon hallinta)
NLP	Natural language processing (luonnollisen kielen käsittely)
PBA	Predictive Business Analytics (ennustava liiketoiminta-analytiikka)
RPA	Robotic Process Automation (ohjelmistorobotiikka)
XML	Extensible Markup Language (merkintäkielen standardi)

# 1 Johdanto

## 1.1 Tutkielman tausta

Schmarzo ja Partlow (2013, s. 1) kuvaavat, kuinka elämme maailmassa, jossa on meneillään dataperusteisen liiketoiminnan vallankumous. He jatkavat, kuinka uuden datan avulla voidaan mahdollistaa yritysten arvonluontiprosessien uudistus. Uuden datan lähteiksi he nimeävät esimerkiksi lukuisat sosiaalisen median lähteet, mobiilit sekä sensori- ja konepohjaisen datan. ”Organisaatiot hukkuvat dataan, mutta samaan aikaan kärsivät tiedon puutteesta” (Maisel & Cokins, 2014, s. 64, Suomennos N.K.).

Davenport (2006, s. 100) korostaa, että elämme aikaa, jona yritysten kilpailukyky on kirstynyt: samalla toimialalla toimivien yritysten tuotteet ja teknologiat ovat hyvin samankaltaisia. Hän kertoo, että hinnalla kilpailu ei länsimaissa onnistu halpamaiden alhaisempien tuotekustannusten vuoksi. Niinpä hän tuo esiin data-analytiikan avulla optimoidut liiketoimintaprosessit ja nimeää ne yhdeksi harvoista alueista, joilla yritykset voivat nykyään erottautua ja kilpailla toisiaan vastaan.

Maailman digitalisoituessa yrityksille avautuu mahdollisuus hyödyntää dataa liiketoiminnassa: kehittää uusia palveluita tai jopa uutta liiketoimintaa (Digia, 2021). Data-analytiikka siis mahdollistaa yrityksen kannattavuuden parantamisen. Haasteena tälle Digian (2021) väitteen mukaan on, kuinka valtavasta datamäärästä saadaan valikoitua tarpeellinen tieto. Tärkeiksi elementeiksi he näkevät tarkoituksenmukaisen datan hyödyntämisen sekä tiedon jalostamisprosessin (keräys, varastointi ja hallinta). Data-analytiikan avulla siis pyritään vastaamaan liiketoiminnalle asetettuihin tavoitteisiin. Data-analytiikan voidaankin ajatella kuuluvan, osastosta riippumatta, jokaisen johtamisalueen asialistalle (Viitala & Jylhä, 2019, s. 138).

Dataa voidaan siis jalostaa ja sen kautta hyödyntää data-analytiikkaa apuna käyttäen. Pärjätäkseen alati kiristyneessä kilpailussa, eri tahot ovat alkaneet hyödyntämään dataa

päätöksenteon tukena. Myös alan aikaisemmissa tutkimuksissa painotetaan datan olennaisuutta ja sen merkitystä liiketoimintaan, päätöksentekoon sekä strategiaan. Siksi onkin mielenkiintoista suorittaa tutkielma, mitä dataa kohdeyrityksissä käytetään hyödyksi ja miten se toimii liiketoiminnan päätöksenteon tukena.

## 1.2 Tutkielman aihe ja tavoitteet

Sekä Business intelligence, että data-analytiikka ovat aiheita, jotka ovat olleet tällä vuosituhannella pinnalla, niin tieteellisessä tutkimuksessa, kuin myös käytännön kentällä. Tieteellisessä tutkimuksessa käsitteitä on tutkittu paljon ja erilaisista näkökulmista. Yksi merkittävistä amerikkalaisista uranuurtajista tieteellisen tutkimuksen saralla on Thomas Davenport, jonka tutkimuksiin tässä tutkielmassa tullaan viittaamaan. Suomalaisessa tutkimuksessa aineistoa on suhteellisen vähän, näistä uusimpana mainittakoon Nykänen ja muut (2016) sekä yhtenä merkittävämpänä Pirttimäki (2007).

Sekä BI, että data-analytiikka ovat käsitteitä, joiden voidaan havaita muuttaneen merkitystään ajan myötä. Tähän on useita syitä, mutta suurin vaikutus lienee digitaalisella vallankumouksella ja sen myötä nopeasti muuttuvalla kehityksellä esimerkiksi datan saavutettavuuden osalta. Siksi onkin tärkeää, että akateemiselle yhteisölle tuotetaan tasaisesti uutta ja ajankohtaista tutkimusta nykytilanteesta.

Suomalaisessa tutkimuksessa kohdeaineistona on toiminut lähinnä kyselytutkimuksen kautta koko Suomen laajuinen vastaajajoukko (Nykänen ja muut, 2016; Pirttimäki, 2007). Näiden molempien voidaan olettaa keskittyneen ja saaneen enemmän vastauksia Etelä-Suomen alueelta, jossa käsitteet ovat yleisempiä. Näiden pohjalta voidaan myös esittää kysymys siitä, mikä tilanne on tällöin muualla Suomessa, esimerkiksi Etelä-Pohjanmaalla? Ennakoasenteena voidaan johtaa ajatus, että pienemmällä alueella data-analytiikan kehitys voi olla vielä eteläistä Suomea jäljessä. Koska vastaavaa tutkimusta ei pelkästään Etelä-Pohjalaisia yrityksiä aineistona käyttäen ole tehty, tämän osalta voitiin katsoa aiheessa olevan sopiva tutkimusaukko.

Tämän empiirisen tutkielman tavoitteena on tuottaa tiedeyhteisölle ajantasainen kuva siitä, käyttävätkö yritykset Business intelligenceä (BI) tai data-analytiikkaa liiketoimintansa päätöksenteon tukena. Tutkielmassa kartoitetaan ensin tilannetta ja selvitetään mistä yritykset hakevat dataa ja mitä työkaluja he käyttävät sen analysoinnissa. Varsinaisena tavoitteena on selvittää, millä tasolla yritykset Etelä-Pohjanmaalla kokevat olevansa BI:n ja data-analytiikan suhteen ja näkevätkö he nykyisen tason riittävänä. Tutkielmassa selvitetään myös yritysten tulevaisuuden näkymiä ja kehittämistarpeita data-analytiikan saralla.

Tutkimuskysymyksiä tutkielmassa on kolme:

- *Mistä tiedonlähteistä päätöksenteon tueksi kerättävä data koostuu ja minkä avulla sitä analysoidaan?*
- *Onko tällä hetkellä saatavissa oleva data ja sen analysointi olennaista ja riittävää liiketoiminnan päätöksenteon tueksi?*
- *Kaivataanko yrityksissä laajempaa data-analytiikkaa, ja mitkä ovat yritysten tulevaisuuden näkymät data-analytiikan suhteen?*

Näiden kolmen tutkimuskysymyksen avulla odotetaan saatavan vastaukset tämän tutkielman tavoitteisiin. Tutkimuskysymysten pohjalta rakennetaan tutkielman teoreettinen viitekehys, jossa käsitellään tutkielman kannalta tärkeimpiä käsitteitä sekä olennaisimpia tieteellisiä tutkimuksia.

### **1.3 Tutkielman rajaukset**

Vaikka tämän tutkielman pääkäsitteet ovat tieteellisessä tutkimuksessa yleisesti tuttuja, yritysmaailmassa ne ovat kuitenkin suhteellisen uusia eikä vastaavaa tutkimusnäkökulmaa vielä löydy. Tieteellisiä tutkimuksia data-analytiikan tilasta suomessa ei juurikaan löydy, joten siltä osin tausta-aineisto nojaa ulkomaiseen tutkimukseen. Suomalaista

aineistoa hyödynnetään soveltuvin osin esimerkiksi BI-tutkimuksien esittelyssä. Vaikka data-analytiikkatason käytön tilaa ei alan akateemisessa tutkimuksessa ole tuotu esiin, tilanne käytännössä yritysten liiketoiminnassa, kuten vertailtaessa saman toimialan kilpailijoihin, on hyvin pinnalla oleva aihe.

Tutkielman aineiston alueeksi rajataan Etelä-Pohjanmaa, koska tutkijan kiinnostuksen kohteena on tarkastella aihetta oman asuinmaakuntansa näkökulmasta. Lisäksi tällä alueella ei ole aikaisemmin toteutettu vastaavaa tutkimusta, joten sen toteutukselle voidaan todeta tarve.

Aikaisemmista tutkimuksista löytyy useita käsitteitä, joilla kaikilla pyrittiin kuvaamaan analytiikkaa. Nämä käsitteet rajataan tässä tutkielmassa kahteen: BI:en sekä data-analytiikkaan. Jälkimmäinen toimii edellisen yläkäsitteenä ja sisältää myös useita muita termejä. Jako näihin kahteen käsitteeseen tehtiin, koska uskottiin, että tutkimusvaiheessa tultaisiin törmäämään BI:en terminä, eikä sitä haluttu sekoittaa data-analytiikkaan. Kuten edellä tullaan esittämään, BI voidaan tulkita varhaiseksi analytiikaksi, kun taas data-analytiikalla viitataan edistyneempään analytiikkaan.

## **1.4 Tutkielman rakenne**

Tämä tutkielma rakentuu kuudesta pääluvusta. Luvut ovat jaettavissa johdantoon, teoriaan sekä empiiriseen osuuteen. Ensimmäinen luku toimii johdantona, joka tutustuttaa lukijan aiheeseen, sen merkitykseen, tavoitteisiin sekä tutkimuskysymyksiin, esittelee rajukset sekä lopuksi tutkielman rakenteen.

Toinen, ja samalla ensimmäinen teoreettisen viitekehyksen aloittava luku, esittelee tutkielman käsitteitä. Tulevien vaiheiden sekä kokonaisuuden ymmärtämiseksi teoriaosuudessa päätettiin esitellä aluksi niin BI:n kuin data-analytiikan tärkeintä raaka-ainetta, dataa. Datan käsitteen alla selvitetään alaluvuissa sen eri muotoja ja ominaisuuksia sekä roolia päätöksenteossa. Käsitteistön esittely jatkuu BI:n määrittelyllä ja päättyy data-

analytiikan esittelyyn. Tutkimusnäkökulman esiin tuomiseksi, data-analytiikan alaluissa vedetään yhteen BI:n ja data-analytiikan erot sekä esitellään data-analytiikan työkalujen roolia.

Kolmannessa luvussa jatketaan teoreettisen viitekehyksen esiin tuomista, keskittyen data-analytiikan käyttöön yrityksissä. Luvussa käsitellään liiketoiminnan haasteita sekä data-analytiikan roolia niissä. Luku jatkuu data-analytiikan haasteiden esittämisellä sekä pohtimalla sen roolia yritysten päätöksenteossa. Niiden jälkeen luku etenee data-analytiikan tulevaisuuden näkymiin. Aikaisempia tutkimuksia aiheista esitellään teorialukujen yhteydessä, jolloin asiayhteys on helpommin tunnistettavissa. Teorialukujen lopuksi toteutetaan tutkielman teoreettinen tarkastelu.

Neljännessä luvussa avataan teoria tutkimusosuuden toteuttamisesta. Luku lähtee liikkeelle tutkimusmenetelmien valinnalla. Se jatkuu aineiston hankinnan suunnitteluun sekä aineiston keräämisen toteutukseen ja päättyy tutkielman aineiston analysointimenetelmien esittelyyn.

Viides luku avaa tutkielman merkittävimmän vaiheen eli tulokset. Luvun alle on jaoteltu aiheet alaluvuittain tutkimuskysymysten mukaisessa järjestyksessä. Luku aloitetaan taustatietojen ja lähtökohtien esittelyllä, josta se etenee datalähteiden ja analysointityökalujen purkamiseen. Näiden jälkeen tuodaan esiin datan sekä analysoinnin olennaisuus ja riittävyys. Aihetta lähestytään avaamalla niin datan ominaisuuksia, kuin liiketoiminnan, että data-analytiikankin, haasteita. Tämän jälkeen luku jatkuu data-analytiikan tämänhetkisen tilan ja riittävyyden sekä tulevaisuuden näkymien esiin tuomisella. Luvun lopuksi esitetään yhteenveto esitetyistä tuloksista.

Viimeisessä, kuudennessa luvussa, tuodaan esiin tutkielmasta tehtävät johtopäätökset. Tutkielmasta annetaan yhteenveto, pohditaan niin tulosten kuin rajoitteidenkin merkitystä sekä esitetään tutkielmasta johdettuja jatkotutkimusehdotuksia.

## 2 Analytiikan käsitteet

### 2.1 Analytiikka

Teoreettisen viitekehyksen avaava ensimmäinen luku lähtee liikkeelle erilaisista käsitteistä, jotka ovat keskeisiä BI:n ja data-analytiikan käsittelyssä, sekä tämän tutkielman omaksumisen kannalta. Alkuun esitellään lyhyesti analytiikka. Toisessa alaluvussa käsitellään dataa, koska lukijalle on tärkeää selvittää, mikä nykypäivänä toimii perustana BI:lle ja data-analytiikalle. Datan alaluvuissa paneudutaan sen laatuun, luotettavuuteen ja rooliin päätöksenteossa. Myös Big data tuodaan esiin omassa alaluvussa, koska sen rooli data-analytiikassa voidaan esittää merkittävänä. Käsitteiden esittely jatkuu tutkielman pääkäsitteiden BI:n sekä data-analytiikan avaamisella. Data-analytiikan alaluvuissa pääkäsitteiden eroavaisuudet tuodaan ilmi ja lisäksi kuvataan työkaluja, joiden avulla dataa valjastetaan toimimaan päätöksenteon tukena.

Watson (2015, s. 33) kuvaa, kuinka analytiikka on tänä päivänä kuuma puheenaihe. Huolimatta sen asemasta nykypäivänä, hän mainitsee sen pitkän historian: analytiikan juuret ulottuvat jopa 1960-luvun loppuun, jolloin ilmestyivät ensimmäiset päätöksenteon tukena toimivat järjestelmät. Maisel ja Cokins (2014, s. 63, 65) huomauttavat, että aiemmin analytiikan käyttö oli harvinaista, mutta nykypäivänä siitä on muodostunut yritysten kilpailuetu. Heidän mielestään analytiikan tehtävänä on käsitellä iso määrä dataa ja tiivistää se käyttökelpoiseksi tietolähteeksi: tällöin analytiikan avulla yksinkertaistetaan dataa ja vahvistetaan sen arvoa.

Analytiikan käsitteistö on laaja ja sen on havaittu muuttuvan riippuen alan nopeasta muutoksesta sekä käsitteistöä käyttävästä tahosta. Nuopponen (2020, s. 111) ilmaisee, että aikaisempiin tutkimuksiin tutustuttaessa esiin nousee niissä käytetty erilainen terministö. Hän neuvoo analysoimaan esitettyä käsitteistöä ja arvioimaan vastaavatko käsitteet sitä, mitä omassa tutkielmassa halutaan tuoda esille. Hän myös näkee teoreettisen viitekehyksen punaisen langan muodostuvan käsitteistön ympärille. Käsitteistön

merkittävyydestä johtuen sen tutkimiseen, ymmärtämiseen ja esittämiseen tulee panostaa. Tässä tutkielmassa on pyritty selventämään pääkäsitteet, kuin myös ajan saatossa muuttunut näkemys käsitteille, jotta ne eivät olisi ristiriidassa nykytiedon kanssa.

## 2.2 Data

Mitä ylipäänsä on data? Pelkälle datalle on vaikea löytää yhtenäistä määritelmää, koska sen käyttö riippuu täysin kontekstista. Tähän tutkielmaan sopiva ilmaisu on esimerkiksi Sorayalla (2018, Glossary), joka määrittelee datan sisältävän faktaa, havaintoja ja raakaa tietoa. Olennaista on, että ilman jatkokäsittelyä hän toteaa datan olevan merkitykse-  
töntä. Tällä määritelmällä voidaan kuvata yhtä hyvin myös raakadataa. Sharda ja muut (2018, s. 57) listaavat datan tärkeimmäksi elementiksi, jota BI ja data-analytiikka tarvitsevat toimiakseen ja organisaatioissa se luetaankin yhdeksi arvostetuimmaksi ominaisuudeksi.

Salo (2013, s. 22, 25) jakaa datan kahteen päätyyppiin: strukturoituun ja strukturoimattomaan dataan. Niiden välimuotoa hän kutsuu semistrukturoiduksi dataksi. Gandomi ja Haider (2015, s. 137) arvioivat, että suurin osa datasta, noin 95 %, on strukturoimattonta dataa. He kertovat sen olevan dataa, joka on saatavilla esimerkiksi äänenä, kuvina, videoina (videomuodossa olevan datan määrä on kooltaan suurinta), sosiaalisessa mediassa tai strukturoimattomana tekstinä. Strukturoitua dataa on yleensä yrityksen itse tuottama data, jota hallitaan esimerkiksi omassa toiminnanohjausjärjestelmässä (ERP) tai asiakkuuksien hallintajärjestelmässä (CRM) (Soraya, 2018, s. 24). Semistrukturoitua dataa taas ovat esimerkiksi merkintäkielen standardi (Extensible Markup Language, XML) -dokumentit, jotka ovat koneellisesti luettavissa niiden sisältämän käyttäjää yksilöivien tunnisteiden vuoksi (Gandomi & Haider, 2015, s. 138). Kuviossa 1 kuvataan datan eri päätyypit, niiden muodot sekä annetaan esimerkkejä dataa sisältävistä kohteista.

<b>Data</b>	<b>Tietomalli</b>	<b>Esimerkit</b>
Strukturoitu	Relaatiotietojärjestelmä	Tietokannat
Semistrukturoitu	XML, CVS, logit	Verkko, logit jne.
Strukturoimaton	Teksti, kuva, video	Sähköpostit jne.

**Kuvio 1.** Strukturoitu, semistrukturoitu ja strukturoimaton data. (Mukaillen Soraya, 2018, s. 25).

Kwon ja muut (2014, s. 388) jakavat datan yrityksen sisäiseen ja ulkoiseen dataan. Näistä sisäiseksi dataksi he luokittelevat datan, joka muodostuu yrityksen liiketoiminnasta, kuten: työntekijöitä, tuotteita tai palveluita, tuotantolinjaa, hallinnollisia päätöksiä, asiakkaita ja heidän ominaisuuksiaan sekä maksuliikennettä tai yrityksen muita resursseja koskevat tiedot. Yrityksen ulkoiselle datalla tunnuksenomaista Kwonin ja muiden (2014, s. 388) mukaan on, että yritys itse ei voi vaikuttaa sen muodostumiseen. Ulkoiseksi dataksi he luettelevat markkina- ja kilpailijatiedot, makrotalouden, muun ulkoisen asiakastiedon sekä toimialan yritystiedon. Ulkoisen datan lähteitä taas ovat esimerkiksi: sosiaalinen media, kännykät, nettisivut ja datasensorit (Bayrak, 2015, s. 231).

Nykypäivänä asiakkaiden luomaa dataa sosiaalisessa mediassa sekä sosiaalisten verkkosivustojen keräämänä muodostuu nopeasti ja sitä on alettu hyödyntämään (Griffin, 2012). Nykäsen ja muiden (2016, s. 40) tutkimuksessa selviää, että ulkoisen, strukturoidun datan käyttö on jo arkipäiväistynyt. Heidän tuloksissaan ilmenee myös, että noin 50 % suomalaisista yrityksistä käyttävät myös sekä sisäistä, että ulkoista semistrukturoitua dataa hyödykseen. Aikaisemmista tutkimuksista voimme päätellä, että strukturoimattomalla, yrityksen ulkoisella datalla on oikein käytettynä iso potentiaali tuoda yrityksille kilpailukykyä. Tällöin potentiaalista uutta dataa voi olla juuri sosiaalisen median kautta tuleva data tai laitteiden sensoreiden tuottama data. Tämän strukturoimattoman, ulkoisen, datan kasvu, on mahdollistanut liiketoiminnalle data-analytiikan hyödyntämisen laajemmin. Vaikka suurin osa datasta on strukturoimatonta ja sitä on myös määrällisesti eniten, organisaatiot eivät osaa hyödyntää sitä tarpeeksi.

Sharda ja muut (2018, s. 59) painottavat datan käsittelyprosessin tärkeyttä ja tähdentävät sen olevan yksi data-analytiikan kriittisimpiä vaiheita. He kiteyttävät, että vaikka organisaatiossa osattaisiin tunnistaa itse ongelma, johon datan avulla haetaan vastausta, saatu vastaus voi olla väärä tai epätarkka, jos datan käsittelyprosessiin ei ole panostettu tarpeeksi. He asettavatkin datalle listan vaatimuksia. Yrityksen toiminnasta riippuu, mitkä näistä vaatimuksista datan tulee täyttää ja mitkä taas voidaan sivuuttaa tai sallia epätäydellisinä. Sharda ja muut (2018, s. 59–60, Suomennos N.K.) listaavat datan vaatimuksiksi:

- lähteen luotettavuuden
- sisällön tarkkuuden
- saavutettavuuden
- tietoturvan ja -suojan
- rikkouden
- johdonmukaisuuden
- ajantasaisuuden
- alkuperäisyyden
- oikeellisuuden
- merkityksellisyyden

Datan määrä on lähtenyt eksponentiaaliseen kasvuun tämän vuosituhannen aikana ja verkossa olevan datan määrä tuplaantuu joka vuosi (Lindholm 2021). Tämä avaa myös yrityksille lukemattomia mahdollisuuksia hyödyntää dataa tai Big dataa.

### **2.2.1 Big Data**

Usein, kun datan yhteydessä puhutaan analytiikasta, datan käsite saattaa muuttua massadatakseksi (Big Data, BD). BD esitetäänkin tässä tutkielmassa data-analytiikan alakäsitteenä. ”Big data ja Big data-analytiikka -termejä käytetään kuvailtaessa valtavan kokoisia ja monimutkaisia tietomääriä ja niiden analyttisiä menetelmiä, jotka vaativat

ainutlaatuista ja edistyksellistä varastointia, hallinnointia, analysointia ja visualisointin menetelmiä” (Chen ja muut, 2012, s. 1166, Suomennos N.K.). Selkeää käännekohtaa siihen milloin data muuttuu BD:ksi ei ole. McKinsey Global Institute (2011) määrittää BD:n muodostuvan datasta, joka on kooltaan niin suurta, ettei sitä voi käsitellä tavallisten ohjelmistojen avulla. Tällöin he näkevät BD:n olevan käsitteenä riippuvainen niin teknologiakehityksestä, toimialasta kuin käytettävästä ohjelmistotyökalusta eikä sen kokoa voida mitata tavuina.

Big dataa lähestytään englanninkielisessä aineistossa usein V-kirjaimen avulla (Salo, 2013, s. 21). BD:a on pyritty määrittelemään useilla eri tavoilla ja käsitteillä. Näistä yksi tunnetuimpia on alkuperäinen Gartnerin (2016) määritelmä, joka kuvaa BD:a kolmen V:n määritelmällä: BD sisältää valtavan määrän (volume), nopeasti uudistuvaa (velocity) ja monimuotoista (variety) tietoa. Soraya (2018, s. 13) kuvaa BD:n kasvua, kuinka tänä päivänä voidaan kerätä tsettatavuja ( $10^{21}$ ) tietoa, kun vielä 10 vuotta sitten puhuttiin megatavuista ( $10^6$ ). Hän myös kertoo BD:n uudistuvan nopeasti, kun joka sekunti uutta dataa muodostuu laitteiden sensorien, sosiaalisen median, sähköpostin ja Googlen kautta. Monimuotoisuudella hän taas viittaa erilaisiin datalähteisiin ja datan eri muotoihin.

Gandomi ja Haider (2015, s. 139) tuovat näiden kolmen V:n lisäksi kolme uutta V:tä määrittelemään BD:a. Totuudenmukaisuudella (veracity), viitataan tiedon luotettavuuteen (Gamage, 2016). Esimerkiksi asiakkaiden näkemykset sosiaalisessa mediassa ovat luonnostaan epävarmaa ja epätarkkaa tietoa, mutta niitä tulee hyödyntää niiden sisältämän organisaatiolle arvokkaan tiedon vuoksi (Gandomi & Haider, 2015, s. 139). Voidaan todeta, että sosiaalisen median käytön ja sähköisyyden kasvamisen vuoksi on johdonmukaista, että totuudenmukaisuus on myöhemmin lisätty mukaan BD:n määritelmään. Gandomi ja Haider (2015, s. 139) esittävät vielä viimeisiksi V:iksi vaihtelevuuden ja monimutkaisuuden (variability and complexity) sekä arvon (value). Vaihtelevuudella he viittaavat datan virtausnopeuksien vaihteluun, monimutkaisuudella taas siihen kuinka BD muodostuu lukemattomien lähteiden kautta. Arvon he tarkoittavat määrittämään BD:n arvoa kuvailevasti. Kuviossa 2 on esitelty BD:n merkittävimmät 6 V:tä esimerkkeineen.



**Kuvio 2.** Big datan 6 V:tä. (Mukaillen Kumar & Thiagarasu, 2019, s. 3).

Lisäksi ajan saatossa V:itä on ilmestynyt lisääkin. Esimerkiksi Soraya (2018, s. 23) esittää listaan vielä 4 V:tä lisää: oikeellisuus (validity), alttius (vulnerability), volatilitteetti (volatility) ja visualisointi (visualization). Dataa käsittelevästä teoriasta voidaan huomata, että sekä datalle, että BD:lle asetetut vaatimukset, ovat osaksi yhtenevät.

Mirarab ja muut (2019, s. 2792) valaisevat, että alkujaan BD:n käytön laajenemisen suurimpina ongelmina on ollut sen kerääminen ja varastoiminen. BD:n varastointi ei

kuitenkaan ole enää nykypäivänä ongelma, koska varastointikustannukset ovat laskeneet roimasti. Janssen ja muut (2017, s. 344) viittaavat muiden V:iden hallinnan olevan nykyään haasteellisempaa, kuin määrän. Acito ja Khatri (2014, s. 567) näkivät varastointikustannusten alenemisen mahdollistaneen data-analytiikan kehityksen. Varastointikustannuksissa on tapahtunut suuri hinnanalennus tämän vuosisadan alussa. Kun datan vaatima tallennustila ja sen kustannukset eivät ole enää rajoittavina tekijöinä, on teknologiatekniikka kehittynyt mahdollisuuden BD:n laajemmalle käytölle. Mirarab ja muut (2019, 2792) tähdentävät, että nykyään monimuotoisuus on BD:n suurimpia haasteita. He nimeävät datamassojen yhdisteltävyyden tärkeäksi elementiksi arvonluomisprosessissa, ja nimeävät keskenään erilaisten datalähteiden hankaloittavan yhdistelyä.

Internet of Thingsin (IoT) valtavirtaistuminen on myös yksi merkittävä tekijä sille, miksi BD:n määrä on kasvanut ja sen käyttö on noussut ilmiöksi (Salo, 2013, s. 12). Sharda ja muut (2018, s. 420) kuvaavat IoT:n olevan yksi nopeimmin kasvavista IT-alan trendeistä. Markkula ja Syväniemi (2015, s. 22, 39) kertovat IoT:n tarkoittavan esineiden ja asioiden internetiä. IoT:n perusideana he tähdentävät, että se mahdollistaa BD-massan muodostumisen tietoverkossa ja sen keräämisen koneiden ja laitteiden, kuten sensorien keskinäisen kommunikoinnin avulla. He siis nimeävät antureiden mittaaman tiedon BD:ksi. IoT:n käyttö on viime vuosina kasvanut räjähdysmäisesti (Mazzuto & Ciarapica, 2019, s. 2). IoT pitää sisällään mahdollisuuksia, jotka ovat vaikuttaneet sen nopeaan kasvuun. Sharda ja muut (2018, s. 420) listaavat näistä kolme merkittävintä: laitteisto, kuten sensorit, ovat muuttuneet halvemmiksi, tehokkaammiksi ja pienentyneet kooltaan, BD ja BI-työkalut ovat helpommin saatavilla ja uusia IoT:n käyttökohteita nousee esiin päivittäin. Koska BD on kasvanut strukturoimattoman datan johdosta, IoT:n merkitys kasvussa on ollut olennainen.

BD:n yhteydessä mainitaan usein myös ydintieto (Master Data, MD) sekä ydintiedon hallinta (Master Data Management, MDM) (Laihonen ja muut, 2013, s. 20). Väre (2019, s. 16) kuvaa MD:n olevan organisaatioissa kokonaisvaltaisesti käytettävää, suhteellisen muuttumatonta dataa, joka kerätään liiketoiminnan ytimestä. Markkulan ja Syväniemen

(2015, s. 39) määritelmän mukaan, MD:llä tarkoitetaan yrityksen liiketoiminnan kannalta kriittistä tietoa, jolle on tunnuksenomaista pysyä lähes muuttumattomana. He listaavat MD:tä olevan esimerkiksi yrityksen asiakas-, tuote- ja henkilöstötiedot. Ydintiedon hallinnalla he kuvaavat MD:n hallinnointia ja kokonaisarkkitehtuuria sisältäen muun muassa laadun ja sisällön valvonnan, tietoturvakysymykset, prosessit ja vastuasiat. Siinä missä MD:n katsotaan sisältävän yrityksen omaa dataa, BD taas sisältää mitä tahansa muuta ulkoista dataa, kuten osto- ja klikkausdataa, sensoridataa sekä some- ja tutkimustietoa (Markkula & Syväniemi 2015, s. 39). Vertailuna MD:hen Väre (2019, s. 18) kirjoittaa BD:n olevan tapahtumatietoa, jota syntyy organisaation perusdatan ulkopuolella. Voidaan siis todeta, että yrityksen datamassan keräys ja hallinnointi lähtee liikkeelle MD:stä, jota tarpeen mukaan laajennetaan BD:llä.

### **2.2.2 Datan laatu ja luotettavuus**

Laihonen ja muut (2013, s. 19–20) tarkastelevat organisaatioiden sisäistä dataa ja painottavat sen olevan usein heikkolaatuista. Ratkaisuksi tähän he ehdottavat, että organisaatioissa keskitytään MD:n puhdistamiseen ja laadun tarkastamiseen. Kwonin ja muiden (2014, s. 391–392) tutkimuksessa ilmenee, että yrityksissä tehdään enemmän dataan pohjautuvia päätöksiä, kun datan laatuun voidaan luottaa. Heidän tuloksissaan merkitystä ei ole siinä, käytetäänkö ulkoista vai sisäistä dataa. Janssen ja muut (2017) tarkastelevat tutkimuksessaan, onko BD:n laadulla vaikutusta sen hyötyyn. He selvensivät, että BD:n avulla tehtävien päätösten laatu riippuu paljolti siitä, minkälainen panostus BD:n prosessointiin on käytetty. Prosessointiin vaikuttavat heidän mukaansa useat tekijät, kuten: BD:n lähteiden ominaisuudet, analysointiprosessin laatu, analysointia tekevien henkilöiden kyvyt, BD:n saatavuus sekä päätöksentekijän ymmärrys datasta ja yhteistyö muiden BD toimintoketjun tahojen kanssa.

Kironin ja muiden (2012, s. 52) tutkimuksessa suositaan datasuuntautunutta yrityskultuuria. Heidän tutkimuksensa osoittaa, että myös innovaatioiden määrä yrityksissä kasvaa silloin, kun päätöksentekijät voivat luottaa datalähteisiin ja niiden prosessointiin.

Watson ja Wixom (2007, s. 96) toteavat datan hankintaosuuden olevan haastavaa ja niimeävät sen yhdeksi kompastuskiveksi datan huonon laadun. Rikhardsson ja Yigitbasioglu (2018, s. 45) totesivat laatuongelmia ilmenevän varsinkin raakadatan käsittelyssä.

Tutkimusten perusteella datan ja sen prosessoinnin laatu vaikuttaisi olevan yksi suurimmista sen käytön haasteista. Shollon ja Galliersin (2016, s. 361) tutkimuksessa datan laatu esitetään ongelmallisena ja sen koetaan vaikuttavan datan luotettavuuteen päätöksenteossa. Toisaalta heidän tutkimuksensa nostaa ilmi, että luotettavuus korostuu enemmän yksittäisissä päätöksissä: pidemmällä aikavälillä tapahtuneet useat havainnot johtavat samansuuntaiseen näkyvyyteen ja tällöin liiketoiminnallinen päätös on mahdollista tehdä, ilman sen perustumista laadukkaaseen dataan.

Kuten edeltävät ja monet muut tutkimukset osoittavat, datan laatu on ollut keskustelunaiheena puhuttaessa datasta, BD:sta tai päätöksenteosta. Datan laadun tutkimuksia haettiin esitellä teoriassa, koska se on aihe, josta myös kenttätutkimuksessa voidaan odottaa kehkeytyvän keskustelua.

### **2.2.3 Data päätöksenteossa**

Päätöksentekoa voidaan harjoittaa eri tavoilla. Kautta aikain päätöksenteko on ollut perusteltavissa mututuntumalla, jonka voidaan katsoa koostuvan päätöksentekijän ammatillisesta osaamisesta ja kokemuksesta. Tällainen päätöksentekoprosessi ei ole perustunut faktatietoon, vaan kirjaimellisesti tuntumaan siitä, mikä voisi olla oikea päätös. Nykypäivänä, kun liiketoiminnan prosesseja viilataan joka suunnalta ja tehokkuushyötyjä haetaan pienistäkin käytännöistä, tällainen päätöksentekomalli ei ole enää perusteltavissa. Datapohjainen päätöksenteko on se, mihin yritykset haluavat perustaa päätöksensä.

Soraya (2018, s. 28) tiivistää datapohjaisen päätöksentekoprosessin vaiheet seuraavasti: ensin päätetään mitä dataa tarvitaan, sitten se kerätään, varastoidaan, esitetään ja

muutetaan tietämykseksi, jonka avulla tehdään päätös. Voidaan ajatella, että varsinkin päätöksentekoprosessin alkupää eli päätös siitä mitä dataa tarvitaan, on merkittävässä roolissa. Seuraavat vaiheet voidaan mieltää datan prosessoinnin vaiheeksi ja viimeisenä vaiheena tulee dataperusteinen päätöksenteko.

Laihonen ja muut (2013, s. 18–19) kertovat, kuinka data itsessään on pelkkää lukumäärärettä. He kirjoittavat, kuinka se sekoitetaan usein tiedon muihin tasoihin. He jatkavat, että data muuttuu informaatioksi vasta, kun se saa rakenteen, ja tietämykseksi (*knowledge*), kun siitä voidaan johtaa analyttistä tulkintaa. Kironin ja muiden (2012, s. 48) tutkimuksessa todetaan, että suurin osa yrityksistä käyttää alle puolet heille käytettävissä olevasta datasta. Tutkimus ei kuitenkaan kerro, sisältääkö käyttämättä jäänyt data hyödyntämätöntä potentiaalia tai toisaalta sitäkään, onko käytettykään data tuonut arvoa tai ollut tärkeää päätöksenteon kannalta.

Khade (2016, s. 986) otaksuu, että nykypäivän suurimpia ongelmia on se, kuinka kaiken saatavilla olevan datan ja eri lähteiden joukosta löytää juuri sen, joka on yritykselle olennaista. Myös Markkula ja Syväniemi (2015, s. 37) kuvaavat yrityksille tärkeäksi voimavaraksi erottaa olennainen data massasta. He toteavat arvon syntyvän silloin, kun data saadaan jalostettua liiketoiminnan ja päätöksenteon kannalta olennaiseksi informaatioksi.

Kirjanpitolaki (1336/1997) määrää paljon olennaisuudesta: olennaisuusperiaatteesta ja olennaisuudesta tilinpäätöksessä. Samassa yhteydessä esitetään myös oikean ja riittävän kuvan antaminen. Olennaisuudelle ei voida määrittää mitään tietynlaista rajaa, vaan tiedon olennaisuus ja sen riittävyys tulee punnita aina tapauskohtaisesti. Kirjanpitolain säännökset toki perustuvat tiettyihin organisaatioiden toimintaa ohjaaviin säännöksiin. Mitä tulee liiketoiminnan päätöksenteossa käytettävän datan olennaisuuteen ja riittävyyteen yrityksissä, rajaa ei vedä mikään ulkopuolinen toimija, vaan jokainen organisaatio määrittää omat säännöksensä, minkä puitteissa datan olennaisuutta ja riittävyyttä arvioidaan.

Datan oikea käyttö päätöksenteon tukena lähtee siitä, että yrityksissä osataan ensin kysyä oikeat kysymykset. Koska tietoa on saatavilla niin paljon, yrityksen tulee olla selvillä siitä, mikä tieto on sen päätöksenteon kannalta olennaista. Kysymykset eivät voi perustua vain dataan, joka on saatavilla, varsinkin kun todellinen arvo toimia päätöksenteon tukena piilee useimmiten ulkoisessa, strukturoimattomassa datassa, joka ei välttämättä ole suoraan saatavilla. Kironin ja muiden (2012, s. 48) tutkimuksessa merkittävä datan käyttöä rajoittava tekijä onkin, että päätöksentekoon tarvittava data ei ole aina saatavilla.

Päätöksentekoprosessi nojaa pitkälti sen takana oleviin ihmisiin. Oikeiden kysymysten esittäjältä vaaditaan syvällistä ymmärrystä yrityksen liiketoiminnasta. On siis selvää, että oikeiden ihmisten avulla yrityksille avautuu lukemattomia mahdollisuuksia kehittää tietoon pohjautuvaa päätöksentekoa. Visio siitä, mitä dataa tarvitaan ja miksi, tulee olla kirkas. Myös laadukas dataprosessi tarvitsee taakseen datanhallinnan asiantuntijuutta. Salo (2013, s. 94) uskoo, että suurin osa BD:sta on tarpeetonta ja vain murto-osalla siitä voidaan tuottaa arvoa yhdistämällä se liiketoimintadataan. Sen löytämiseksi hän kertoo tarvittavan teknologiaa, mutta korostaa ihmisillä ja osaamisella olevan prosessissa merkittävin rooli.

Päätöksentekijöiden tulee voida luottaa siihen, että heille kerätty tieto on luotettavaa. Jos data ei ole luotettavaa, se voi organisaatiossa aiheuttaa intuitiopohjaisen päätöksenteon kasvun. Kowalczykin ja Buxmannin (2015, s. 10) mukaan intuitiopohjainen päätöksenteko, varsinkin ei-rutiininomaisissa päätöksissä, muodostaa organisaatiossa riskin. Kun päätöksentekoprosessi on kunnossa, vastaukset kysymyksiin saavutetaan ja datasta saadaan irti parhain hyöty. Tällöin yritys voi saavuttaa merkittävää etua kilpailijoihinsa nähden.

## 2.3 Business intelligence

Tässä luvussa esitellään ensimmäinen tämän tutkielman pääkäsite. Kyseessä on liiketoimintatiedon hallinta, joka myös varsinkin yritysmaailmassa tunnetaan paremmin Business intelligençenä, BI:nä. Englanninkielisen termin, kuin myös siitä johdetun lyhenteen, voidaan katsoa vakiintuneen yritysmaailmassa, ja niitä tullaan käyttämään myös tässä tutkielmassa.

BI:n englanninkielinen termi muuttui yleiseksi ja ilmestyi päätöksentekijöiden sanastoon 1990-luvun alussa (Chen ja muut, 2012, s. 1166; Watson, 2015, s. 33). Vuonna 2006 ilmestyivät kaupalliset BI-termiä kantavat tuotteet ja palvelut (Sharda ja muut, 2018, s. 15). Termin merkityksen voidaan todeta muuttuneen kehityksen kautta. Chen ja muut (2012, s. 1166–1168) sekä Olszak (2016, s. 107) jakavat BI:n kehityksen kolmeen eri aikaan: BI 1.0, BI 2.0 ja BI 3.0. He uskovat termin saaneen alkunsa jo 1970-luvulla, jolloin hyödynnettiin pääosin sen ajan järjestelmistä saatavaa strukturoitua dataa. Ennen vuosituhaten taitetta heidän mukaansa siirryttiin vaiheeseen BI 2.0., jolloin organisaatiot alkoivat käyttää myös strukturoimatonta verkkopohjaista dataa. BI 3.0 vaiheen murroksessa vuonna 2005 mukaan tulivat kännykkä- ja sensoripohjainen data (Chen ja muut, 2012, s. 1168–1169; Olszak, 2016, s. 107).

BI-termillä ei voida katsoa olevan yksiselitteistä merkitystä ja se voidaan tulkita eri tavoin käyttäjästä riippuen. Käytännön näkökulma voi olla hyvinkin erilainen, jos kyseessä on it-asiantuntija, ohjelmantoinittaja, liiketoiminnan asiantuntija tai muu yrityksen loppukäyttäjä. Isik ja muut (2013, s. 13) sekä Olszak (2016, s. 107) arvioivat, että BI:ä voidaan lähestyä joko teknologisesta tai organisaatiollisesta näkökulmasta. Tieteellisessä tutkimuksessa tutkijat ovat pyrkineet tuomaan esiin termin monet merkitykset.

Pirttimäki (2007, s. 57–64) on väitöskirjassaan tunnistanut BI:lle useita määritelmiä. Hän on tulkinnut sen johtamiskäsitteeksi, joka voidaan jakaa kahteen päämääritelmään: jalostettuun tietoon sekä prosessiin. Kolmantena, mutta ei yhtä merkittävänä,

tarkoituksena, hän tunnistaa teknologisen määritelmän. Toisaalta Olszak (2016, 106–107) on listannut BI:lle jopa yhdeksän eri merkitystä:

- Työkalut, teknologiat ja ohjelmistot
- Tiedolla johtaminen (knowledge management)
- Päätöksentekojärjestelmät
- Raportointinäkyvät (dashboards)
- Dataan tai tiettyyn menetelmään suuntautunut kulttuuri
- Prosessit
- Analytiikka
- Älykäs kilpailukyky
- Big data

Kaikki edellä mainitut määritelmät voidaan tunnistaa osaksi BI:ä, mutta tässä tutkielmassa ne voidaan yhdistellä muutamaankin pääkategoriaan: teknologiseen ja prosessinäkökulmaan sekä analytiikkaan.

Sharda ja muut (2018, s. 467) määrittelevät BI:n teknologisesta näkökulmasta, jolloin se sisältää työkaluja, sovelluksia, IT-arkkitehtuurin sekä tietokannat. Watson ja Wixom (2007, s. 96) arvioivat BI-termiä käytettävän liike-elämässä kuvaamaan analyttisiä sovelluksia. Davenport (2006, s. 102) uskoo, että perinteinen BI on IT-yhteisön käyttämä termi prosesseille ja ohjelmistolle. Tarkasteltaessa BI:ä pelkästään ohjelmistonäkökulmasta, se voidaan ajatella työkaluna, joka kerää ja esittää tietoa. Toisaalta Shollo ja Galliers (2016, s. 343) esittävät, että tällä hetkellä hyväksytyin määritelmä BI:lle olisi, että se on yhdistelmä eri näkökulmia: teknologiaa, tuotetta ja prosessia.

Jos BI kuvataan prosessina, tällöin se sisältää teknologianäkökulman. Watson ja Wixom (2007, s. 96–97) kuvaavat prosessin, joka muodostuu datan hankkimisesta ja tietovarastoinnista (*data warehousing*) sekä datan käyttämisestä päätöksenteon tukena. Jälkimmäistä he kutsuvat BI:ksi. BI-prosessin päätarkoituksena on tukea ja tuottaa parempaa

päätöksentekoa (Davenport, 2006, s. 106–107; Isik ja muut, 2013, s. 13–14). Toisaalta Shollo ja Galliers (2016, s. 343, 357) kritisoivat näkemystä, että BI itsessään tuottaisi parempaa päätöksentekoa. He tutkivat aihetta näkökulmasta, jossa BI on osa yrityksen kokonaisprosesseja. Tällöin merkittävämpinä tutkimustuloksina he nostavat esiin kommunikoinnin sekä datan valinnan tärkeyden prosessin onnistumisessa.

Joissain aikaisemmissa tutkimuksissa BI-termin on todettu käsittävän vielä laajemman kokonaisuuden. Davenport ja Harris (2007, s. 32) kertoivat BI:n olevan hyvin laaja alue ja sisältävän myös analytiikan. Tämä vaikuttaisi johtuvan siitä, että data-analytiikan termi ei vielä tuolloin ole ollut vakiintuneessa käytössä. 2010-luvulle tultaessa tieteellisissä tutkimuksissa termi BI tuntuu väistyvän data-analytiikan käsitteen tieltä. Sharda ja muut (2018, s. 22, 130) havaitsivat, kuinka termiä BI käytetään nykyään kuvaamaan data-analytiikan alkuvaiheita, kuten kuvailevaa analytiikkaa. He väittävät, että data-analytiikka on nykypäivänä korvannut tietokonepohjaisista päätöksentekoteknologioista aiemmin käytetyn terminologian, myös termin BI. Watsonin (2015, s. 33) mukaan BI:ä on aikaisemmin käytetty laajempänä terminä, joka on sisältänyt niin teknologioita, prosesseja kuin sovelluksia. Hän väittää, että nykyään data-analytiikka toimii näiden toimintojen kattoterminä. Tässä tutkielmassa BI:llä todetaan olevan kaksi merkitystä: teknologinen sekä prosessi-näkökulma. Analytiikka esitetään BI:n yläkäsitteenä: data-analytiikkana.

Laihonen ja muut (2013, s. 46) kertovat tiedon etsimisen ja merkityksen arvioimisen päätöksenteon tukena olevan BI:ä. He uskovat, että BI:ä käytetään näin joka organisaatiossa jollakin tasolla. He jatkavat, että käyttöä ei vain aina ole tunnistettu ja siksi ei välttämättä osata kertoa BI:n käytöstä. Tämän voidaan tulkita johtuvan siitä, että organisaatioissa BI mielletään vahvasti teknologiapainotteiseksi työkaluksi.

## **2.4 Data-analytiikka**

Tässä luvussa esitellään toinen, ja samalla myös merkittävin, tämän tutkielman pääkäsite. Data-analytiikka käsitteenä on ristiriitainen ja siksi ensimmäisessä alaluvussa

perustellaan sen valintaa sekä soveltuvuutta tutkielmaan. Toisessa alaluvussa data-analytiikkaa tullaan vertaamaan toiseen pääkäsitteeseen, Business intelligenceen, jotta tutkimusasetelma saadaan esiin. Viimeisessä alaluvussa tuodaan esiin data-analytiikassa hyödynnettäviä työkaluja.

#### **2.4.1 Data-analytiikan käsite**

Termejä data-analytiikan ympärillä on paljon. Varsinaisesti termi data-analytiikka ei sellaisenaan ole tieteellisessä tutkimuksessa käytetty termi. Analytiikka taas on tämän tutkielman kannalta liian laeva termi. Maisel ja Cokins (2014, s. 63) huomauttavat, kuinka nykypäivänä analytiikalla on useita määritelmiä. Omassa tutkimuksessaan he vaihtavat analytiikan termiksi ennustava liiketoiminta-analytiikka (predictive business analytics, PBA), koska se kuvastaa paremmin heidän tutkimaansa analytiikan suuntautumista. Data-analytiikan tai pelkän analytiikan lisäksi terminä tässä tutkielmassa toimivat liiketoiminnan analytiikka (Business Analytics, BA), Big Data (BD) tai Big Data Analytiikka (BDA), Liiketoimintatiedon hallinta ja analytiikka (Business Intelligence & Analytics, BI&A), ennustava analytiikka tai PBA. Myös tiedolla johtaminen tai tietojohdaminen (*knowledge management*)(Viitala & Jylhä, 2019, s. 136–137) sisältyvät termeihin. Jotta tämä tutkielma olisi johdonmukaisempi, edellä mainitut termit on yhdistetty ja niistä käytetään yhtä nimitystä: data-analytiikka. Yhteisestä määritelmästä on suljettu pois ainoastaan BI, joka käsitellään omana määritelmänään. Data-analytiikka valikoitui sopivammaksi nimeksi, koska tässä tutkielmassa halutaan painottaa datan merkittävää roolia analytiikassa.

”Data-analytiikka, on laaja kattotermi, jolla tarkoitetaan datan käsittelyä ihmiselle hyödynnettävään muotoon. Analytiikasta puhuttaessa tarkoitetaan usein erityisesti prosessin loppupäätä: datan hyödyntämistä ja tulosten esittelyä ihmisille.” (TietoEvry, 2020).

Data-analytiikan määritelmän voidaan katsoa olevan laaja ja monesti sitä pyritäänkin kuvaamaan kokonaisuutena. Aciton ja Khatrin (2014, s. 566) sekä Chenin ja muiden (2012,

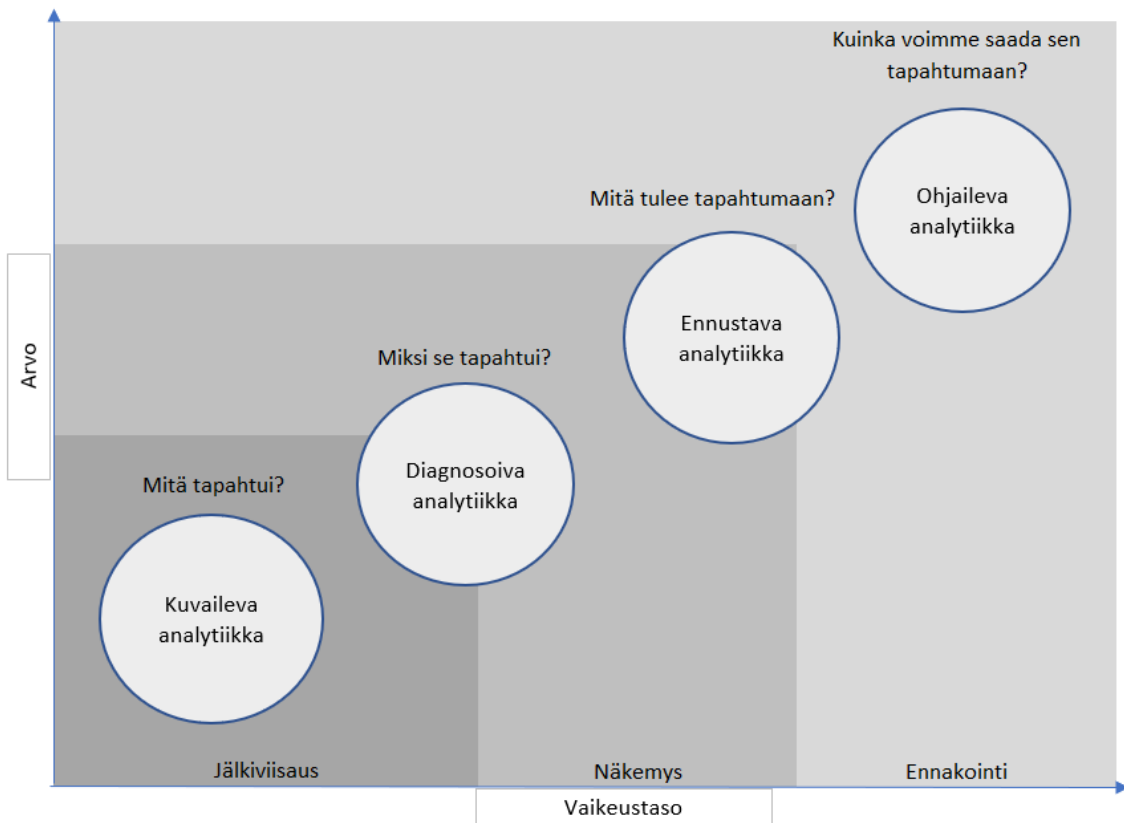
s. 1166) käytännönläheinen näkemys on, että data-analytiikan tavoitteena on tehdä datasta arvokasta yritykselle. He viittaavat termillä datan prosessointiin kokonaisuutena, joten termi sisältää niin teknologian kuin menetelmätkin. Bayrak (2015, s. 231) sekä Maisel ja Cokins (2014, s. 69) nimeävät, että data-analytiikka edistää datan avulla tehtäviä oivalluksia ja sen rooli on parantaa organisaatioiden päätöksentekoa. Oivallukset ovat yrityksille tärkeitä, koska niiden avulla yleensä saadaan liiketoiminnasta uutta tietoa. Oivallusten avulla voidaan tuottaa kilpailuetua, parantaa kannattavuutta tai vaikka kehittää uutta liiketoimintaa. Data-analytiikan roolia organisaatioiden päätöksenteossa käsitellään myöhemmin (ks. luku 3.2 Data-analytiikka päätöksenteossa ja data-analytiikan haasteet).

Osassa tutkimuksista data-analytiikka on erotettu muista vastaavista termeistä edistyneiden ominaisuuksien avulla. Davenport ja Harris (2007, s. 26) kuvaavat data-analytiikan sisältävän tilastollista ja kvantitatiivista analyysia sekä selittäviä ja ennustavia malleja. Olszak (2012, s. 107) viittaa tiedonlouhintaan ja Watson (2015, s. 34) yleisiin matemaattisiin tekniikoihin. Institute for Operations Research and Management Sciences (INFORMS) mukaan ”data-analytiikalla tarkoitetaan tietotekniikkaa, johtamistieteiden tekniikoita ja tilastotieteitä yhdistävää menetelmää, joka ratkaisee todellisia ongelmakohtia” (Sharda ja muut, 2018, s. 22–23, Suomennos N.K.). Nämä määritelmät data-analytiikasta on esitetty eritoten tieteellisestä näkökulmasta.

Sircar (2009, s. 293) jakaa data-analytiikan menetelmät perus- ja ennustavaan tasoon. Hän luettelee optimoinnin, ennustavan mallintamisen, tilastollisten analyysien sekä ennustamisen olevan data-analytiikan menetelmiä, joiden avulla voidaan tuottaa kilpailukykyä. Osassa tutkimuksista data-analytiikkaa on pyritty määrittämään BD:n ja sitä määrittelevien V:iden avulla. Kwon ja muut (2014, s. 387) sekä Begum ja George (2017, s. 1650) kuvailevat data-analytiikan avulla analysoidavan suuria määriä monimutkaista dataa.

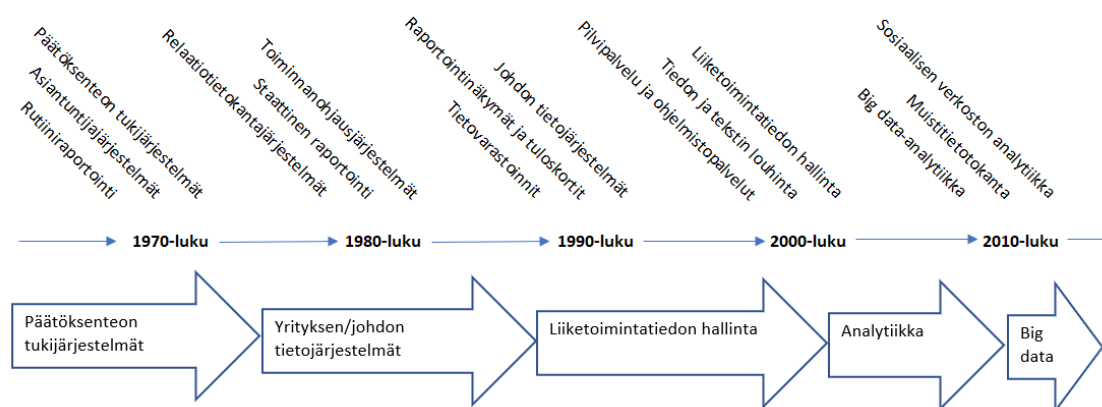
Koska data-analytiikan termin voidaan katsoa käsittävän myös teknologian ja ohjelmistot, jotkut tutkijat ovat käyttäneet tätä lähestymiskulmana. Torres ja muut (2018, s. 822–823) lähestyvät data-analytiikkaa teknologisesta näkökulmasta ja määrittelevät sen keskittyvän analyttisiä tekniikoita tarjoaviin sovelluksiin. Markkula ja Syväniemi (2015, s. 16–17) painottavat termillä viitattavan vahvasti strategiaan, sen jalkauttamiseen sekä seurantaan analysoitua tietoa hyväksikäyttäen.

Appelbaum ja muut (2017, s. 32) esittelevät data-analytiikan kolme eri menetelmää: kuvailevan (descriptive), ennustavan (predictive) ja ohjailevan (prescriptive) analytiikan. He kertovat kuvailevan analytiikan tarvitseman datan koostuvan historiatiedosta, jolloin kuvailevaa analytiikkaa käytetään monesti perustana yrityksen seurantajärjestelmissä. Gandomi ja Haider (2015, s. 143) kertovat ennustavan analytiikan tavoitteena olevan nimensä mukaisesti ennustaa liiketoiminnan tapahtumia. He havaitsevat ennustavassa analytiikassa suuren potentiaalin ja huomauttavat, että sitä voidaan tuottaa joka toimialalla ja lukemattomilla eri tekniikoilla. Appelbaum ja muut (2017, s. 32) kirjoittavat yritysten luovan ohjailevan analytiikan avulla erilaisia vaihtoehtoisia lopputuloksia. Menetelmiin voidaan lukea myös diagnosoiva analytiikka (diagnostic analytics), jonka avulla historiallisesta tiedosta haetaan keskinäistä korrelaatiota (Viitala & Jylhä, 2019, s. 138). Kuviossa 3 on esitetty data-analytiikan eri menetelmät ja niiden haastavuus sekä potentiaalinen arvontuotto suhteessa toisiinsa.



**Kuvio 3.** Data-analytiikan menetelmät. (Mukaillen Gartner, 2019)

Koska tieteellisessä tutkimuksessa ei tunnisteta termiä data-analytiikka, käsitteen syntymisestä ei voida myöskään esittää suoraa faktatietoa. Voidaan esittää, että data-analytiikkaa on ollut olemassa jo pitkään, se vain on esiintynyt eri nimillä ja kehityskaarensaan muodostanut sen data-analytiikan käsitteen, mikä sillä esitetään olevan nykypäivänä ja tässä tutkielmassa. Kuviossa 4 on esitetty data-analytiikan elämäнкаari ja siitä aikaisemmin käytetty termistö.



**Kuvio 4.** Data-analytiikan elämänskaari. (Mukaillen Sharda ja muut, 2018, s. 13).

## 2.4.2 Business intelligenen ja data-analytiikan erot

Edellisissä luvuissa esitetyt määritelmät BI:stä sekä data-analytiikasta ovat osittain samankaltaisia ja päällekkäisiä. Tämä johtuu osin muuttuneesta termistöstä, mutta myös näkökulmasta, kuka ja kuinka termejä tarkastellaan. Siksi onkin hyvä saattaa ilmi, miten akateemisessa tutkimuksessa näiden termien eroja käsitellään.

Sircar (2009, s. 290) rinnastaa BI:n ja data-analytiikan termit tutkiessaan liiketoiminnan opetusta kouluissa. Tällöin hän esittää BI:n olevan IT-yhteisön käyttämä termi, kun taas data-analytiikkaa suositaan liiketoiminnan parissa. Kwon ja muut (2014, s. 390) erottavat BI:ssä käytettävän sisäistä ja virallista, esimerkiksi toiminnanohjausjärjestelmän (ERP), dataa, kun taas data-analytiikassa käytetään enemmän ulkoista ja epävirallista dataa esimerkiksi sosiaalisesta mediasta.

Aiemmin tässä tutkielmassa (ks. luku 2.2 Business intelligence) BI on esitetty teknologisesta ja prosessinäkökulmasta. Rikhardsson ja Yigitbasioglu (2018, s. 38) kuvailevat data-analytiikkaa samoilla määritelmillä, mutta he keskittyvät sen laajaan käyttöön ja nostavat esiin BD:n sekä mobiili- ja sensoridatan käyttämisen osana data-analytiikkaa.

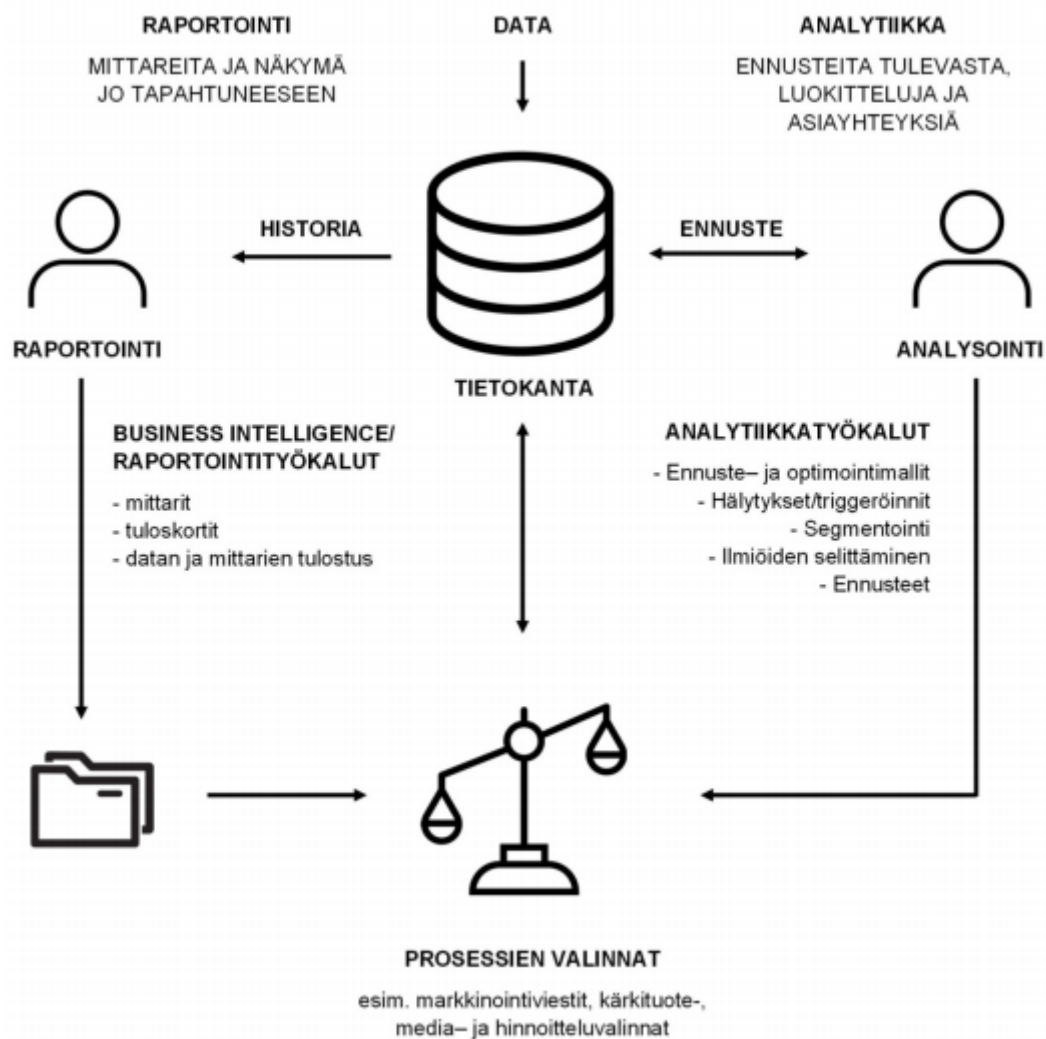
Maisel ja Cokins (2014, s. 65) erottavat BI:n ja data-analytiikan käsitteet toisistaan. He uskovat, että perinteisen BI:n tarkoitus on kerätä ja varastoida historiallista dataa ja käyttää sitä vastaamaan yrityksen perustavanlaatuisiin kysymyksiin. Sen sijaan data-analytiikan he kertovat tarjoavan uutta, eteenpäin katsovaa informaatiota ja kysymyksiin vastaamisen sijaan sen avulla luodaan uusia kysymyksiä. He arvelevat, kuinka monissa yrityksissä BI esitetään lopullisena määränpäänä, sen sijaan että data-analytiikka tulisi esittää sellaisena. He siis kiteyttävät, että BI ei sellaisenaan riitä tuottamaan uusia oivalluksia vaan siihen tarvitaan data-analytiikkaa lisäksi. Huomioitavaa tutkimuksessa on, että BI:ä ja data-analytiikkaa ei esitetä toisiaan poissulkevinä vaihtoehtoina. Tämän vuoksi voidaan tulkita, että data-analytiikka on kokonaisuus, johon tähdätään, ja BI-teknologia on osa data-analytiikkaa.

Termien väliset erot voidaan erotella organisaatiossa myös käytännön ja niiden hallinnan avulla. Davenport (2006, s. 102) toteaa BI:n osastotasaisen hallinnoinnin aiheuttavan ongelmia, kun taas data-analytiikkaa käyttävät tahot luovat keskitettyjä ryhmiä vastaamaan datan hallinnasta. Hän pohtii, että jokaisen osaston erikseen tekemä työ aiheuttaa päällekkäisyyttä ja päällekkäisyyden poistamisella voidaan saavuttaa valtavia kustannus- ja tehokkuushyötyjä.

Davenport ja Harris (2007, s. 26) käsittelevät BI:n isona kokonaisuutena ja esittävät data-analytiikan sisältyvän siihen. He listaavat esimerkiksi optimoinnin ja ennustamisen osaksi BI:ä. Nykypäivänä termit ovat kääntyneet toisinpäin eli BI luetaan pikemminkin osana data-analytiikkaa. Soraya (2018, glossary) totesi data-analytiikan olevan monessa yhteydessä käytetty termi, joka pitää sisällään BI:n ja kaikki analyttisiin sovelluksiin liittyvät toiminnot. BI:n hän kuvailee olevan kattotermi sovellukselle, työkaluille sekä IT-ympäristölle. Voidaankin todeta, että hän on esittänyt BI:n teknologisesta näkökulmasta ja tulkinnut sen olevan data-analytiikan alla toimiva alakäsite.

Markkula ja Syväniemi (2015, s. 89–90) nimeävät molemmat termit edistyneeksi analytiikaksi. He kuitenkin määrittävät BI:n esittävän historiatietoa kuvailevaa

analytiikkaa käyttäen, kun taas data-analytiikalla viitataan tulevaisuuteen ja ennustamiseen, sekä käytetään ennustavaa ja ohjailevaa analytiikkaa. Kuvio 5 havainnollistaa BI:n ja data-analytiikan sekä niiden työkalujen eroja.



**Kuvio 5.** Business intelligenen ja analytiikan erot. Markkula ja Syväniemi (2015, s. 90).

Tässä tutkielmassa termi data-analytiikka toimii kattokäsitteenä. Voidaan todeta, että data-analytiikka on yläkäsite, johon myös BI sisältyy. Termi BI on data-analytiikan alakäsite ja sillä tarkoitetaan joko raportointityökalua, kuten BI-ohjelmistoa, tai raportointia, joka perustuu historiatietoon. Näin ollen data-analytiikalla tarkoitetaan kaikkea muuta käsitteistön ympärillä pyörivää termistöä ja ennen kaikkea esimerkiksi tulevaisuuteen nojaavaa analytiikkaa, kuten ennustavaa analytiikkaa.

Pirttimäki (2007) tutkii väitöskirjassaan BI:n käyttöä suomalaisissa suuryhtiöissä. Hänen tutkimuksensa vahvistaa, että BI:n ei koeta olevan tarpeeksi tehokas ja kattava menetelmä. Syyksi tähän tutkija pohtii, että BI-hankinnoissa on inhimillisen näkökulman sijaan keskitytty liiaksi teknologiaan. Olszak (2016) selvittää tutkimuksessaan, mitä haasteita BI:n käyttöönottoon liittyy. Hän esittää, että monissa organisaatioissa BI:ä ei saada toimimaan tehokkaasti päätöksenteon tukena, saati tekemään sen avulla olennaista kilpailuetua. Merkittävämpänä löydöksenä tutkimuksessa ilmenee, että monet kohdeyrityksistä ovat BI:n kehityksen kanssa vielä aikakaudella BI 2.0. Inhimilliset ja organisaatiolliset tekijät, kuten strategia ja kulttuuri, mielletään hyödyllisimmiksi, kuin teknologia.

Nykänen ja muut (2016, s. 32–33) tutkivat BI:n käyttöä päätöksenteon tukena. Tuloksista ilmenee, että yli puolet vastaajista kokevat huonolaatuisen tai riittämättömän datan olevan syynä BI:n arvon heikentymiseen. Tutkimuksesta ilmenee myös, että vain joka kuudennen vastaajan mukaan BI:ä käytetään yrityksissä joka osastolla.

Vera-Baquero ja muut (2015, s. 215) havaitsivat tutkimuksessaan, että liiketoimintaprosessien parantaminen perinteisten BI-työkalujen avulla ei ole tehokasta. He listasivat haasteita olevan erityisesti yhtiöillä, jotka käsittelevät isoa määrää monimuotoista, strukturoimatonta dataa. BI-työkaluja ei heidän mukaansa ole tehty prosessikeskeisen liiketoiminnan tarpeisiin ja ne eivät kata koko organisaation tarpeita. Liiketoimintaprosesseista he nostavat esiin toimitusketjun ja pohtivat sen tehostamiseen käytettävien edistyneempiä tekniikoita sekä pilvipohjaisia ratkaisuja.

Edellä esitetyt tutkimukset osoittavat erityisesti haasteita, joihin organisaatiot ovat törmänneet BI:n käytössä. Tämän tutkielman oletuksena on, että käyttämällä BI:n sijaan siitä kehittyneempää muotoa, data-analytiikkaa, näihin haasteisiin voitaisiin vastata.

### 2.4.3 Data-analytiikan työkalut

Tieteellisissä tutkimuksissa puhutaan sekä BI- että data-analytiikan järjestelmistä ja työkaluista. Voikin olla vaikea tunnistaa ja jaotella, mitkä ovat niiden erottavat tekijät. Watson ja Wixom (2007, s. 96) toteavat, että BI:ä käytetään, erityisesti yrityksissä, kuvaamaan data-analytiikan työkaluja. BI-järjestelmät voivat olla perinteisiä, jolloin ne nojautuvat historialliseen dataan tai edistyneempiä, jolloin niiden avulla voidaan muodostaa kehittyneempää data-analytiikkaa, kuten ennustamista. Tässä tutkielmassa on aikaisemmissa luvuissa esitetty BI:n olevan osa data-analytiikkaa, joten selkeän tutkimusnäkökulman säilyttämisen vuoksi, myös tässä osiossa voidaan esittää BI-työkalujen olevan osa data-analytiikan työkaluja.

Eritoten käytännön kentällä työkaluista puhutaan siis usein BI-ohjelmistoina tai niistä käytetään jotain sopivampaa, sen käyttötarkoitusta kuvaavaa termiä, kuten ennustamisjärjestelmä. Kuten edellä Watsonin ja Wixomin (2007) tutkimuksesta voidaan johtaa, työkaluista käytettävät termit voivat vaihdella sekä tulkitsijasta, että käyttöympäristöstä johtuen.

Shollo ja Galliers (2016, s. 341–342) kuvaavat, kuinka organisaatioiden käyttämät järjestelmät ovat kehittyneet viimeisen 60 vuoden aikana yrityksen toimintaa ja päätöksentekoa tukevista järjestelmistä BI-järjestelmiin. BI-järjestelmien mahdollisuuksien he uskovat olevan monipuolisessa tietojen hyödyntämisessä, koska ne käyttävät lähteenä dataa, joka on kerätty yrityksen eri tietokannoista. Datan yhdistelemisen mahdollistumisen myötä he näkevät BI-järjestelmien avulla tuotettavan uusia oivalluksia ja sen myötä parempaa päätöksentekoa.

Begum ja George (2017, s. 1651) mainitsevat, että BI-työkaluja käytetään ensisijaisesti yrityksen tietokannoista saatavan datan esittämiseen. He toteavat, kuinka BI-työkalujen avulla käyttäjät voivat helposti luoda raportteja sekä visuaalisia esityksiä. He mieltävät BI-työkalut erityisen tehokkaiksi silloin, kun käyttäjien on tehtävä nopeita päätöksiä ja luotava uusia raportteja päätöksenteon tueksi. Datan muuntaminen visuaaliseen

muotoon on tärkeä osa päätöksentekoprosessia, koska visuaalinen näkymä lisää käyttäjän ymmärrystä tietoon ja sen kautta auttaa tuottamaan uusia oivalluksia.

Sharda ja muut (2018, s. 98–99) sekä Watson ja Wixom (2007, s. 98) lähestyvät työkaluja raportointinäkökulmasta ja esittävät tehokkaimpia visuaalisia menetelmiä olevan: raportointinäkymä (*dashboard*), painotettu tuloskortti (*balanced scorecard*) ja mittarit, kuten suorituskykymittari (*key performance indicator, KPI*).

Sircar (2009, s. 294) tarkastelee data-analytiikan työkaluja ja toteaa niitä olevan hyvin eritasoisia ja erilaisiin käyttöön tarkoitettuja. Hän jakaa työkalut ryhmiin: perinteiset taulukkopohjaiset työkalut, tilastolliset työkalut, edistyneemmät BI-työkalut, toiminnanohjausjärjestelmien yhteyteen luodut analysointiosiot ja teollisuudessa käytettävät ennustavat sovellukset. Begum ja George (2017) taas tutkimuksessaan listaavat data-analytiikan työkaluja, jotka on tarkoitettu erityisesti suuren tietomassan käsittelyyn. He uskovat, että data-analytiikan työkalujen tehtävänä on vastata yrityksen erikoistarpeisiin. Nykäsen ja muiden (2016, s. 40) tutkimuksessa selviää, että taulukkopohjaiset työkalut ovat eniten käytetyin analysointimuoto. Lisäksi heidän tutkimuksessaan tiedon visualisointijärjestelmät nousevat toiseksi käytetyimmäksi työkaluksi.

Eritelläksemme BI- ja data-analytiikan työkalujen ominaisuuksia, voimme päätellä, että perinteiset BI-järjestelmät ovat hyviä työkaluja silloin, kun kyse on perinteisestä historiaan perustuvasta raportoinnista tai päätöksenteosta sekä sen visualisoinnista. Tällöin ne voidaankin nimetä yrityksen perustoimintoja tukeviksi järjestelmiksi. Data-analytiikan työkalut taas ovat mutkikkaampia: ne keräävät dataa ulkoisista lähteistä, data voi olla strukturoimatonta, sen prosessointiin voidaan käyttää edistyksellisiä menetelmiä, kuten tiedonlouhintaa ja ohjailevaa analytiikkaa ja niiden tavoitteena on tuoda yritykselle uutta arvoa.

Watson ja Wixom (2007, s. 98) kertovat, että BI-järjestelmien avulla datan käyttö on reaaliaikaista ja onnistuu mistä ja milloin vain, aikaa säästyy ja käyttäjä voi tehdä nopeita

huomioita sekä ajantasaisia päätöksiä. He lausuvat, ”mitä seurataan, se tehdään” (Suomenkoski N.K.). Toisaalta kiteyttävää sanontaa voidaan tulkita myös toisesta näkökulmasta. Jos käyttäjä seuraa mittareita vain niiltä osin, kuin hänelle niitä raportointinäkymissä tarjotaan, tällöin hallinta siitä, mitä seurataan ja mistä tehdään päätöksiä, on annettu näkymän luoneen tahon käsiin.

Rikhardsson ja Yigitbasioglu (2018, s. 42) kertovat, kuinka edistyneet data-analytiikan työkalut sallivat nykyään käyttäjille mahdollisuuksia itse muokata datan esittämistapaa. Tällä he vertaavat tilannetta perinteiseen tapaan, jossa joko ohjelmistotoimittajat tai talousosasto vastaavat raporttien muokkausmahdollisuuksista. Voidaan tulkita, että perinteiset BI-ohjelmistot eivät välttämättä ole muokattavissa esittämään dataa siinä muodossa kuin käyttäjä niitä tarvitsee. Sharma ja muut (2014, s. 438) painottavat perinteisten BI-ohjelmien olevan vaihtoehto, joka edelsi data-analytiikan työkaluja.

Sharma ja muut (2014, s. 434), Shollo ja Galliers (2016) sekä Nykänen ja muut (2016, s. 34) kritisoivat ajatusta, että työkalut itsessään tuovat organisaatiolle arvoa tai tuottavat parempaa päätöksentekoa. Heidän tutkimuksistaan voidaan johtaa ajatus, että data-analytiikan työkalut toimivat päätöksenteon apuvälineinä, joiden avulla data yhdistetään ja muunnetaan visuaaliseen muotoon. Itsessään ne eivät toimi päätöksenteossa, vaan väliin tarvitaan ihminen tulkitsemaan ja analysoimaan tuotoksia.

Aikaisemmista tutkimuksista ilmenee myös, että data-analytiikan työkalun käyttö yrityksissä ei aina ole perusteltua. Tutkimusten mukaan työkalun tarve lähtee liikkeelle ongelman tunnistamisesta tai strategiasta, mutta yritykset saattavat investoida IT-hankintoihin, koska kilpailijat ovat niin tehneet tai koska se mielletään ajankohtaiseksi trendiksi tai sosiaaliseksi arvonnousuksi (Kwon ja muut, 2014, s. 388; Radaelli, 2000).

### 3 Data-analytiikka yrityksissä

Edeltävässä teorialuvussa esiteltiin tutkielman käsitteistöä. Tässä luvussa pohditaan yleisellä tasolla haasteita, joita yrityksen liiketoiminnassa voi ilmetä sekä tarkastellaan data-analytiikan roolia niissä. Luku jatkuu esittelemällä data-analytiikan roolia organisaation päätöksenteossa sekä data-analytiikan haasteita. Data-analytiikan käytössä on havaittu lukuisia haasteita, joihin data-analytiikkaa harjoittavat toimijat ovat törmänneet, joten tutkielmassa halutaan tuoda esiin myös kriittistä näkemystä data-analytiikan käytöstä. Kolmannessa pääluvussa pohditaan lyhyesti data-analytiikan tulevaisuuden näkymiä sekä merkittävimpiä kehityksen trendejä. Lopuksi esitetään vielä lyhyt yhteenveto tutkielman tavoitteista ja kerrataan tutkielmaa ohjaavat tutkimuskysymykset.

Data-analytiikasta ei ole olemassa valmista mallia, jota voitaisiin tarjota organisaatioiden käyttöön. Jokainen yritys on erilainen, niiden haasteet ovat erilaisia ja päätöksenteko on erilaista. Ei liene alaa, jossa data-analytiikan käytöstä ei olisi hyötyä, mutta on myös toimialoja, joissa sen käytöstä voidaan hyötyä muita enemmän. Davenport ja Harris (2007, s. 25) nimeävät data-analytiikkaa eniten käytettäviksi aloiksi päivittäistavarakaupan, vähittäiskaupan, rahoituspalvelut, hotellit ja ravintolat, lääketeollisuuden ja teollisuuden, kuljetuksen ja logistiikan, internetin ja telekommunikaation. Näiden lisäksi esimerkiksi tilintarkastus on Suomessa ala, jossa data-analytiikan käyttö on ollut pinnalla viime vuosina.

Data-analytiikan käyttöä voidaan lähestyä myös suoraan organisaationäkökulmasta, jolloin sen käyttöä voidaan tarkastella osastoittain. Schmarzo ja Partlow (2013, s. 4–5) luettelevat yrityksen eri osastoja, jotka hyötyvät data-analytiikan käytöstä. He nimeävät hankinnan, tuotekehityksen, tuotannon, logistiikan, markkinoinnin, hinnoittelun, myynnin sekä oheismyynnin, myymälät ja henkilöstöhallinnon tällaisiksi. Voidaankin todeta, että kirjoittajat uskovat data-analytiikan käytön olevan mahdollista organisaation lähes joka osastolla. Näiden lisäksi listaan voisi vielä lisätä esimerkiksi talousosaston sekä verkkokaupan.

### 3.1 Liiketoiminnan haasteet ja data-analytiikka

Jokaisessa liiketoiminnassa toimialasta riippumatta on omat haasteensa, joihin tehokkaalla johtamisella sekä eri menetelmiä käyttäen pyritään vastaamaan. Haaste tulee ensin tunnistaa, jotta siihen voidaan lähteä hakemaan vastausta datan avulla (Davenport, 2013, s. 122).

Kiron ja muut (2012, s. 48) tutkivat, että data-analytiikan käyttö kannustaa uusiin liiketoiminnan innovaatioihin organisaatioissa. Innovaatioiden avulla he havaitsivat saattavan tehokkuushyötyjä eritoten markkinointi- ja talousosastoilla, kuin yleisesti operatiivisessa toiminnassa. Begum ja George (2017, s. 1650) ovat samoilla linjoilla ja mainitsevat asiakaspalvelun ja markkinoinnin tehostamisen, sekä niin operatiivisen työn, kuin kilpailukyvyyn parantumisen, olevan alueita, joihin data-analytiikan avulla voidaan hakea parannusta. Voidaan todeta, että näillä parannuksilla sekä tehokkuushyödyillä organisaatiot voivat saavuttaa liiketoiminnassaan esimerkiksi suurempaa tuottoa tai kokonaisvaltaisesti parempaa palvelua.

Lawless (2014) mainitsee data-analytiikan avulla yrityksissä saavutettavan lukemattomia tehokkuusetuja. Hän keskittyy erityisesti kysynnän ennustamiseen. Tällöin data-analytiikan tuoma lisäarvo huomataan siinä, että ymmärrys asiasta kasvaa ja johtaa parempaan päätöksentekoon. Lisähyötyjä hän esittää useita, jotka johtavat koko liiketoiminnan optimoimiseen. Hän lisää, että kun asiakasymmärrys kasvaa, se johtaa tuotteiden parempaan kohdentamiseen, useampiin liideihin, markkinoinnin ja mainostamisen tehostamiseen, liikevaihdon ja tuoton kasvuun sekä aineellisiin taloudellisiin hyötyihin. Tästä voidaankin huomata, kuinka data-analytiikan avulla voidaan vaikuttaa erityisen paljon myynti- ja markkinointitoimintoihin ja niiden kautta koko yrityksen suorituskykyyn.

Lawless (2014, s. 46) listaa myös liiketoiminnallisia haasteita, joiden ennustamiseen voidaan hakea vastauksia data-analytiikan avulla. Toimialakohtaisina haasteina voidaan luetella esimerkiksi sähkönkulutus, hinnanmuodostus osakemarkkinoilla, etukuponkien käyttöaste, rikollisuus, äänestysprosentti, opintojen keskeyttämisprosentti, elokuvien

suosittelu, tapaturmat, lainahäiriöt ja kuolleisuus. Yleispätevinä haasteina taas mainitaan tuotteiden kysynnän ennakointi, tuotetuntemus ja työssä vaadittavat taidot, henkilöstön vaihtuvuus sekä asiakaspoistumat. Näiden lisäksi haasteina voidaan tuoda esiin esimerkiksi asiakkaiden maksuhäiriöt, toimitusketjun tai tuotantoprosessin hallinta sekä koneiden ja laitteiden vikaantuminen.

Sircar (2009, s. 293) uskoo, että olemalla parempi kuin kilpailija, luomalla uusia liiketoiminnallisia mahdollisuuksia sekä ennakoimalla asiakkaita paremmin, organisaation on mahdollista saavuttaa kilpailukykyä. Tämä vaatii muutoksia organisaation liiketoimintaprosesseihin ja data-analytiikan saralla menetelminä voidaan käyttää niin ennustamista, kuin optimointia. Myös Schmarzo ja Partlow (2013, s. 9) huomauttavat parhaimman ratkaisun etsimisen olevan tavoiteltavaa ja analyttisten sovellusten mahdollistavan automaattisen liiketoimintaprosessien optimoinnin.

Rikhardssonin ja Yigitbasioglun (2018, s. 45) uskovat organisaatioiden saavan vastauksia liiketoiminnan haasteisiinsa sen avulla, kuinka paljon tarkoituksenmukaista dataa ne onnistuivat löytämään. Kwon ja muut (2014, s. 387) taas uskovat päätöksenteon vaikutuksen liiketoiminnalle olevan merkittävämpi data-analytiikan myötä. He kuitenkin toteavat hyötyjen olevan saavutettavissa vain prosessiin vaadittavan oikeanlaisen osaamisen avulla. Nämä tutkimukset herättävät ajatuksen, että haasteiden selättämisen voidaan mieltää riippuvan paljolti dataa keräävien, tai sen keräämisestä päättävien, henkilöiden kyvykkyydestä. Data-analytiikkaa hallitsevien henkilöiden osaaminen voidaan siis nykypäivänä ja eritoten tulevaisuudessa mieltää kriittiseksi data-analytiikan menestystekijäksi.

### **3.2 Data-analytiikka päätöksenteossa ja data-analytiikan haasteet**

Data-analytiikka on nykyään osa päätöksentekoprosessia kaikissa isommissa organisaatioissa (Bayrak, 2015, s. 230). Lönnqvist ja Pirttimäki (2006, s. 32) nimeävät tehokkaan liiketoimintatiedon tuottamisen yrityksille välttämättömäksi elementiksi. Data-

analytiikan avulla kilpailukykyä voidaan luoda ja päätöksentekoa parantaa organisaation joka osastolla. Kwon ja muut (2014, s. 387) listaavat data-analytiikan avulla saavutetuksi hyödyiksi organisaatioissa kustannusten alenemisen sekä uudet liiketoiminnalliset suuntauksset.

Päätöksenteko voidaan jakaa kolmeen eri kategoriaan: strategiseen, taktiseen ja operatiiviseen päätöksentekoon (Bayrak, 2015, s. 230). Maisel ja Cokins (2014, s. 66) nimeävät strategiset päätökset yritystoiminnan kannalta merkittävämmiksi päätöksiksi. Toisaalta he jatkavat, että operatiiviset päätökset ovat päivittäisiä ja niitä tehdään määrältään huomattavasti enemmän, joten niiden tueksi data-analytiikasta saadaan suurin hyöty. Kwon ja muut (2014, s. 392) havaitsevat, että ulkoisista lähteistä kerätyn datan avulla voidaan saavuttaa suurempi hyötyä strategisen päätöksenteon tueksi. Rikhardssonin ja Yigitbasioglun (2018, s. 44) esittävät, että data-analytiikkaa voidaan käyttää kaikissa eri päätöksentekokategorioissa, mutta itse päätöksentekoprosessi saattaa usein olla hajainen.

Viitala ja Jylhä (2019, s. 136) kertovat, että datasuuntautunut päätöksenteko ja sen laatu on alettu nostamaan osaksi yritysten strategiaa. Davenport ja muut (2001) nostavat esiin strategian ja sen tärkeyden prosessissa, muutettaessa dataa tietämykseksi. Heidän tutkimuksessaan strategia kuvataan datasuuntautuneen organisaation perustana. He havaitsevat strategian selkeyttävän analyttistä prosessia ja kirkastavan organisaation tavoitteita. Päämäärän tavoittelu voidaan esittää olevan strategian perimmäisenä tavoitteena (Viitala & Jylhä 2019, s. 62). Strategian avulla data-analytiikka voidaan liittää osaksi organisaation liiketoimintaprosesseja, jolloin data saadaan tukemaan päämäärää.

Davenport (2006) tutkii yrityksiä, jotka ovat omaksuneet data-analytiikan osaksi strategiaansa. Hän havaitsee käytäntöjä, jotka tutkimuksessa osoittautuvat yhteneviksi data-analytiikkaa hyödyllisimmin käyttävien yritysten osalta. Kolme tärkeintä havaintoa ovat: data-analytiikkaa lähestytään liiketoiminnallisesta näkökulmasta ja sitä sovelletaan koko yrityksen tasolla, johtavassa asemassa olevilla henkilöillä on data-analytiikkamyönteinen

asenne ja he johtavat esimerkillään, kolmantena mallinnusta ja optimointia, kuten ennakkoivaa mallinnusta, sekä ulkoista dataa, käytetään näissä yrityksissä laajasti. Näiden havaintojen lisäksi tutkimuksessa tunnistetaan vielä neljä muuta menestystekijää: oikea kulttuuri, jolloin päätöksenteko pohjautuu aina faktalle, oikeat ihmiset, jotka ovat erikoistuneet analyyttiseen ajatteluun, oikea teknologia, jonka avulla organisaatioon luodaan toimiva datastrategia, sopiva BI-ohjelmisto sekä riittävän tehokas laitekanta sekä oikea fokusointi, jolloin päätöksentekoprosessi on linjassa kokonaisstrategian kanssa ja resurssit osataan suunnata sitä eniten palveleville toiminnoille. Hän myös havaitsee, että onnistunut data-analytiikan omaksuminen saattaa organisaatioissa tarkoittaa vuosien työtä. Hän nimeää, että liiketoimintaprosessien yhtenäistäminen datastrategian kanssa ja teknologiset osa-alueet, kuten lähtödatan putsaus ja ohjelmiston valinta ja käyttöönotto ovat aikaa vieviä kokonaisuuksia. Muutos vaatii myös rahallisesti merkittäviä investointeja.

Rikhardsson ja Yigitbasioglu (2018, s. 44–45) uskovat data-analytiikkapohjaisen päätöksenteon nojautuvan datassa ilmeneviin toistuvuuksiin ja kaavamaisuuksiin. Tällöin päätöksenteko heidän mukaansa perustuu kaavaan, ei niinkään syyllä, joka sen aiheuttaa. He jatkavat, että päätöksenteko voi tapahtua hyvinkin nopeasti, kun reaaliaikaisen datan avulla eri vaihtoehdot ja niiden vertailu voidaan suorittaa nopeasti. He nimeävät myös haasteita datakeskeiseen päätöksentekoon siirtymisessä. Tutkimuksessaan he mainitsevat erityisiksi ongelmakohtiksi erilaiset analysointitavat, datalähteet sekä menetelmät.

Davenport ja muut (2001, s. 125–128) nimeävät kriittiseksi menestystekijäksi data-analytiikan saralla sen, kuinka yrityskulttuuri saadaan muutettua arvostamaan dataperusteista päätöksentekoa. He jatkavat suurimman roolin tämän osalta olevan organisaation johdolla ja heidän asenteellaan datapohjaista päätöksentekoa kohtaan. Myös Markkula ja Syväniemi (2015, s. 17) nimeävät yrityksen rakenteiden muuttamisen olevan ensimmäinen liiketoiminnallinen haaste data-analytiikkaan kääntyessä. Näiden tutkimusten myötä voidaan olettaa, että mitä isompi organisaatio on kyseessä, sitä syvempään uurretut ovat sen prosessit ja organisaatorakenne. Vanhojen rakenteiden purkaminen

ja uusien toimintamallien omaksuminen ovat iso haaste, lähdetessä suuntaamaan kohti data-analytiikan omaksumista.

Myös Maisel ja Cokins (2014, s. 69) esittävät analyyttisen yrityskulttuurin luomisen sekä faktaperusteisen päätöksenteon yrityksen menestystekijöinä. Vastauksena näiden luomiselle he mainitsevat muutoshallinnan sekä käyttäytymismuutoksen nostamisen ylemmäs organisaation tärkeysjärjestyksessä. Bhatt (2001, s. 74–75) tuo yrityskulttuurin luomiseksi esiin teknologian, henkilöstön ja menetelmät kaikki yhtä tärkeinä elementteinä ja esittää, että ratkaisu on saada niiden keskinäinen suhde tasapainoiseksi. Voidaan todeta, että tutkimuksissa on korostettu henkilöstön roolia data-analytiikan omaksumisessa. Sharma ja muut (2014, s. 435, Suomennos N.K.) toteavatkin, että ”oivallukset syntyvät analyytikoiden ja liiketoiminnan johtajien välisestä aktiivisesta prosessista”. He jatkavat, että vaikka organisaatio käyttäisi kuinka edistyksellisiä data-analytiikan menetelmiä, prosessissa on aina jossain vaiheessa ihminen päättämässä asioista tai hyväksymässä niitä.

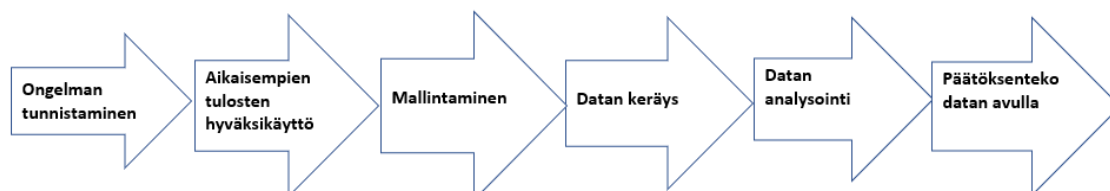
Janssen ja muut (2017) tutkivat data-analytiikkaan pohjautuvan päätöksenteon laatuun vaikuttavia tekijöitä. He huomauttavat, että data-analytiikan käyttö ei automaattisesti tarkoita parempaa päätösten laatua ja että datan ymmärtäminen ja sen arvon esiin saattaminen ovat onnistumisen suurin haaste. Samaa kriittisyyttä esittävät myös Sharma ja muut (2014, s. 433–434), jotka eivät otaksu arvontuottoa automaattiseksi hyödyksi. He jatkavat, kuinka hyötyä saavutetaan jo siinä, että liiketoimintaprosessit uudistetaan mahdollistamaan data-analytiikan käyttö. Tämän jälkeen itse data-analytiikan käytön osuus organisaation suorituskyvyn parantumisessa tuo lisätehokkuutta liiketoimintaan.

Sharda ja muut (2018, s. 379) kiteyttävät haasteet ja esittivät kaikkiaan seitsemän eri tekijää, jotka tulisi ottaa huomioon, jos data-analytiikasta halutaan saada parhain hyöty. Näitä olivat:

- Selkeä liiketoiminnan tarve

- Edistyneen analytiikan taitoja omaava henkilöstö
- Oikeat analyttiset työkalut
- Vahva datainfrastruktuuri
- Tietoon perustuva päätöksentekokulttuuri
- Yhtenäinen linjaus liiketoiminnan ja IT:n välillä
- Vahva ja asialleen omistautuneen johdon tuki

Davenport (2013, s. 122) esittää data-analytiikkapohjaisen päätöksenteon jakamista kuuteen vaiheeseen: Päätöksentekoprosessi lähtee liikkeelle ongelman tunnistamisesta tai kysymyksen esittämisestä, johon datan avulla haetaan vastausta. Viimeiseen vaiheeseen kuuluu tulosten esittely ja toimiminen eli päätöksenteko datan avulla. Päätöksentekoprosessin pilkkominen osiin tuo esiin mielenkiintoisen havainnon. Kun koko prosessi lähtee liikkeelle ongelman tunnistamisesta, tällöin kaikki seuraavat toimet tukevat sitä, miten ongelmaan saadaan vastaus. Kuviossa 6 havainnollistetaan data-analytiikkapohjaisen päätöksentekoprosessin etenemistä.



**Kuvio 6.** Data-analytiikkapohjainen päätöksentekoprosessi.

Myös Lönnqvist ja Pirttimäki (2006, s. 37) jakavat analytiikkaprosessin vastaaviin vaiheisiin. Heidän jaottelunsa sisältää neljä vaihetta, jotka ovat: tietotarpeen yksilöinti, tiedon hankinta, analysointi ja varastointi sekä tiedon hyödyntäminen. Kuten huomataan, molempien tutkimusten prosessit alkavat ja päättyvät samoihin teemoihin.

### 3.3 Data-analytiikan tulevaisuuden näkymät

Davenport (2018, s. 74–76, 79) tuo esiin tulevaisuuden data-analytiikan suuntauksia. Hänen mukaansa data-analytiikan uusi aikakausi toimii jatkumona kehitykselle, joka edistyksekkisissä organisaatioissa on saavutettu kokonaisvaltaisessa data-analytiikan käytössä. Hän lukee tulevaisuuden data-analytiikkaan suuresti kuuluvaksi tekniikaksi tekoälyn (*Artificial Intelligence, AI*), kuin myös kone- ja syväoppimisen (*Machine and deep learning*), ohjelmistorobotiikan (*Robotic process automation, RPA*) sekä luonnollisen kielen käsittelyn (*Natural language processing, NLP*). Tämän uuden aikakauden hän uskoo, erityisesti AI:n käyttöpotentiaalin myötä, olevan yksi mullistavimmista teknologisista muutoksista koskaan.

Viitala ja Jylhä (2019, s. 182–183) viittaavat tekoälyllä algoritmeihin perustuviin ohjelmistoihin, joiden päätöksenteko sekä oppiminen on verrattavissa ihmisajatteluun. Ohjelmistorobotiikalla he kuvaavat automatisoituja prosesseja, joissa ohjelmisto hoitaa tiedon keruun ja yhdistelyn ihmisen sijasta. Davenport (2018, s. 75) kertoo perinteisen koneoppimisen olleen jo hyvän aikaa käytössä yrityksissä. Tällä hän viittaa koneoppimisen hyödyntämiseen ennustavassa analytiikassa. Tulevaisuuden koneoppimisella hän kuitenkin viittaa monitahoisempiin algoritmeihin ja menetelmiin, kuten päätöspuumalliyhdistelmiin. Hyvinkin monimutkaista syväoppimisen neuroverkkoa, hän taas kertoo käytettävän erityisesti ääni- ja kuvatunnistuksessa.

Vaikka yleisesti edellä mainittujen käsitteiden osalta voidaan puhua tulevaisuuden suuntauksista, joitain tekoälyn asteita on jo käytössä yrityksissä. Tässä tutkielmassa kuitenkin oletuksena on, että tarkastelualueella edistyneen tekoälyn käyttö on vielä vähäistä ja laaja hyödyntäminen osa tulevaisuutta.

### 3.4 Teoreettinen tarkastelu ja tutkimuskysymykset

Tämän tutkielman tavoitteena on selvittää, millä tasolla BI:n tai data-analytiikan käyttö ovat Etelä-Pohjanmaalla. Kuten edellisissä luvuissa on osoitettu, käsitteiden välillä voidaan tehdä useita huomioita. BI:llä voidaan vastata yrityksen perustavaa laatua oleviin kysymyksiin, kun taas data-analytiikalla tähdätään lisäarvoon. BI:n avulla jalostetaan enemmän sisäistä dataa, kun taas data-analytiikkaa käyttävät panostavat enemmän ulkoiseen dataan, jolla voidaan saavuttaa yritykselle merkittävämpää hyötyä. BI voidaan aikaisempien tutkimusten valossa esittää nykypäivänä olevan osa data-analytiikkaa ja data-analytiikan olevan BI:stä seuraava kehittyneempi muoto. Molempien käyttö pohjautuu vahvasti niiden perusteena käytettävään dataan, sen laatuun ja tarkoituksenmukaisuuteen, kuin teknologiaan. Siksi onkin mielenkiintoista tutkia, millä tasolla tilanne yrityksissä Etelä-Pohjanmaalla tällä hetkellä on.

Tutkimuskysymyksinä tutkielmassa toimivat:

- *Mistä tiedonlähteistä päätöksenteon tueksi kerättävä data koostuu ja minkä avulla sitä analysoidaan?*
- *Onko tällä hetkellä saatavissa oleva data ja sen analysointi olennaista ja riittävää liiketoiminnan päätöksenteon tueksi?*
- *Kaivataanko yrityksissä laajempaa data-analytiikkaa, ja mitkä ovat yritysten tulevaisuuden näkymät data-analytiikan suhteen?*

Ensimmäisen kysymyksen tehtävänä on kartoittaa yrityksen lähtötilanne ja päätöksenteon raaka-aine. Tarkentavien kysymysten avulla tulisi ensinnäkin selvittää käytetäänkö yrityksessä ulkoista vai sisäistä ja strukturoitua vai strukturoimatonta dataa. Toiseksi kysytään, miten data päätyy päätöksentekijän tarkasteluun, eli mitä työkaluja sen jalostamiseen ja esittämiseen käytetään.

Toisella kysymyksellä haetaan vastauksia useampiin teemoihin. Ensinnä keskitytään raaka-aineen ominaisuuksiin sekä teknologiaan ja kysytään, onko datan laatu ja työkalut riittäviä. Toiseksi haastateltavan organisaatiosta kerätään lisätietoa ja selvitetään

minkälaisia haasteita he liiketoiminnassaan näkevät. Tätä aihetta syvennetään ja pohditaan, käytetäänkö päätöksenteossa siihen sopivaa dataa ja työkaluja. Kolmantena teemanäkökulmana kysytään olennaisuudesta, eli uskotaanko yrityksessä olevan tarpeetonta raportointia, joka ei palvele päätöksentekoa.

Kolmannessa kysymyksessä kartoitetaan tyytyväisyyttä nykytilanteeseen sekä pohditaan tulevaisuutta. Nykytilanteesta haetaan tuntemusta, koetaanko prosessi riittäväksi ja millä tasolla yritys on data-analytiikan käytössä. Tulevaisuudesta pohditaan, onko data-analytiikkaa tarve laajentaa ja onko sitä kartoitettu.

## 4 Tutkielman toteutus

Ennen varsinaisen kenttätutkimuksen osuutta tuodaan esiin, kuinka empiriaosuus tullaan toteuttamaan. Tässä luvussa kuvataan tulevaa tutkielman empiiristä osuutta ensinnäkin menetelmällisestä näkökulmasta, sen jälkeen kerrotaan aineiston hankintamenetelmistä ja kuinka ne käytännön tasolla toteutetaan. Lopuksi tuodaan esiin aineiston analysointimenetelmät.

### 4.1 Tutkimusmetodologia

Tässä tutkielmassa käytettiin tutkimusmenetelmänä kvalitatiivista tutkimusta ja tarkemmin sen alakäsitettä, vertailevaa tapaustutkimusta. Alasuutari (2012) omaksuu oikean tutkimusmetodin määrittelemisen tärkeäksi osaksi laadukasta tutkimusta. Hänen mukaansa tutkimusmetodi luodaan, jotta tutkimuksesta voidaan muodostaa havaintoja ja jotta tutkija voi erottaa omat havaintonsa tutkimuksen tuloksista. Hän myös jatkaa, että tutkimusmetodi pitää sisällään rajauksia, joiden sisällä tutkija voi toimia. Ominaiseksi oikealle tutkimusmetodille hän nimeää yhteisen linjan teoreettisen viitekehyksen kanssa.

Puusa ja muut (2020) ilmaisevat tutkittavien henkilöiden näkökulman tavoittamisen olevan ominaista kvalitatiiviselle tutkimukselle. He määrittelevät laadullisen tutkimuksen hyödyt siinä, että tämän menetelmän avulla tutkimuksessa voi syntyä jotain ennalta-arvaamattomia tuloksia. Kvalitatiivinen tutkimus sopiikin tämän tutkielman metodiksi siksi, että ennalta ei ole tiedossa minkälaisia tuloksia empiirisessä osiossa on odotettavissa.

Puusa ja muut (2020) kirjoittavat vertailevan tapaustutkimuksen ominaispiirteistä. Heidän mukaansa tämän menetelmän vuoksi tutkimuksen osissa voidaan todeta joko samankaltaisuutta tai niitä voidaan erottaa toisistaan. Tämä menetelmä sopi kyseessä olevaan tutkielmaan parhaiten, koska tuloksissa voitiin olettaa olevan odotettavissa samankaltaisuutta, huolimatta siitä, että ne toimivat esimerkiksi erilaisilla toimialoilla.

## 4.2 Aineiston hankinta

Puusan ja muiden (2020) mukaan haastattelut ovat useimmiten käytetty metodi, kun puhutaan laadullisen tutkimuksen toteutuksesta. He arvioivat haastattelututkimuksen eduksi sen, että tutkijalla on vapaus valita haastattelujen kohteeksi henkilöitä, joilla oletettavasti on kokemusta käsiteltävästä aiheesta. He jatkavat, että tällöin on mahdollista saada haastatteluista irti tietoa kattavasti ja haastattelutilanteessa henkilöille voidaan esittää tarkentavia kysymyksiä. Koska kyseessä oli pienimuotoinen laadullinen tutkielma, haastattelujen perusteella ei voitu tehdä yleistystä tuloksista. Tarkoituksena olikin kerätä ja tulkita otosta aiheesta.

Tämä tutkielma toteutettiin haastattelututkimuksilla. Hirsjärvi ja Hurme (2008, s. 35) esittävät haastattelututkimuksen olevan hyvä vaihtoehto myös silloin, kun kyseessä on vähän kartoitettu alue, ja tutkija ei voi etukäteen arvioida, millaisia vastauksia on tulossa. He myös mainitsevat haastattelujen hyväksi puoleksi tutkijan mahdollisuuden syventää tietojaan haastattelutilanteessa.

Haastateltavien yritysten joukon kartoitus tehtiin käyttäen Orbis-tietokantaa. Kartoitus suoritettiin käyttäen rajauksena Etelä-Pohjanmaata kotipaikkanaan pitäviä aktiivisia osakeyhtiöitä, joiden liikevaihto (tilikausilla 2019 tai 2020) oli yli 100 miljoonaa euroa. Tiedot tarkistettiin käyttäen apuna asiakastiedon yrityshakua asiakastieto.fi. Tutkimuksen kohteeksi valideja yrityksiä löytyi yhdeksän kappaletta.

Tämän tutkielman tavoitteena on antaa kokonaiskuva alueella toimivien ennalta rajattujen yritysten data-analytiikan tilasta. Yksittäisten yritysten osalta data-analytiikan tilaa ei tuoda esiin. Tulokset perustuvat haastateltavien henkilöiden näkemukseen aiheesta, mutta henkilöitä on haastattelutilanteessa pyydetty vastaamaan kysymyksiin edustamansa yhtiön näkökulmasta. Voidaan myös todeta, että lähtökohtaisesti henkilöiden työympäristö, sen vakiintuneet tavat sekä vallitsevat prosessit ohjaavat henkilöiden näkemystä. Näiden seikkojen vuoksi haastateltavien henkilöiden nimet sekä heidän

edustamiensa yhtiöiden nimet päätettiin salata. Haastattelujen saamisen uskottiin myös helpottuvan salaamisesta johtuen.

Aikaisempiin tutkimuksiin perehtyminen osoitti, että kokonaisvaltaisen data-analytiikan ymmärtämiseen ja hallintaan vaaditaan laaja-alaista liiketoimintaosaamista. Tästä syystä yrityksistä valittiin haastatteluihin henkilöitä, jotka toimivat monipuolisesti yrityksen liiketoiminnan tehtävissä. Yrityksissä lähestyttiin johto- ja päällikköasemassa olevia henkilöitä, joilla uskottiin olevan monipuolista näkemystä edustamansa yrityksen data-analytiikasta ja päätöksenteosta. Kaikkia yhdeksää validia yritystä lähestyttiin haastattelupyynnöillä. Tavoitetuista henkilöistä yhteensä kuusi kappaletta suostui antamaan haastattelun tutkielmaa varten. Määrän voitiin katsoa olevan riittävä tutkimusjoukkoon verrattuna.

Tämä tutkielman aineiston hankinta suoritettiin puolistrukturoidulla teemahaastattelulla. Koska etukäteen ei voitu tietää, mitä tuloksia tutkielmassa saadaan, koettiin paremmaksi antaa haastateltaville vapaus kertoa omin sanoin ja näkemysin ilmiöstä, ilman että tutkijan oma näkemys tilanteesta vaikuttaa haastateltavaan. Tutkimuskysymykset, jotka toimivat myös teemahaastattelukysymyksinä, toimitettiin etukäteen haastateltavien tiedoksi. Haastattelutilanteessa haastateltaville esitettiin tarkentavia kysymyksiä ilmiöstä, heidän antamistaan näkemyksistä riippuen. Haastattelukysymykset voidaankin tämän perusteella jakaa pääkysymyksiin sekä tarkentaviin kysymyksiin.

Haastattelutyylejä tunnetaan useita erilaisia, joista tähän tutkielmaan valittiin sen luonteeseen parhaiten sopivat menetelmät. Hirsjärven ja Hurmeen (2008, s. 47–48) mukaan puolistrukturoitujen haastattelujen tyypillinen piirre on, että haastattelujen jotkin näkökulmat, kuten kysymykset, on ennalta päätetty, mutta vastaukset on jätetty vastaajan vapaasti päätettäväksi. He nimeävät teemahaastattelun eduiksi, että haastattelussa tuodaan haastateltavan tulkinta paremmin esiin ja haastateltavan antama merkitys ilmiölle nousee keskeiseen osaan.

Kohdeyritysten sähköiseen materiaaliin tutustuttiin ennakkoon. Yrityksiin tutustuminen perustui pitkälti heidän nettisivuiltaan löytyneeseen tietoon, kuten vuosikertomuksiin ja muuhun sähköiseen aineistoon, kuten lehtiartikkeleihin. Tutustuminen tehtiin, jotta haastattelutilanteessa olisi parempi valmius ymmärtää haastateltavan lähtökohtia. Samasta syystä tutustuttiin myös haastateltavien henkilöiden työtaustaan, johon apuna käytettiin lähinnä LinkedIn-sivustolta löytyvää sosiaalisen median materiaalia.

Tutkielman aineistona käytettiin Suomessa yksilohaastatteluilla kerättyä aineistoa, jonka kerääminen suoritettiin elo-syyskuun 2021 aikana. Haastatteluja suoritettiin yksi kappale per kohdeyritys. Teemahaastattelukysymykset toimitettiin haastateltaville ennakkoon, jotta kohdehenkilöillä oli aikaa tutustua haastattelun sisältöön ja tarvittaessa selvittää asioiden tilaa oman työnantajayrityksensä osalta.

### **4.3 Aineiston analysointi**

Puusa ja muut (2020) kertovat aineiston analysoinnin tavoitteena olevan tuottaa hyvin perusteltua tulkintaa käsiteltävästä aiheesta sekä tehdä oivalluksia sen perusteella. He esittävät, että haastatteluaineisto itsessään on haastateltavien luomaa subjektista raakadataa käsiteltävästä aiheesta. He jatkavat, että tällöin tutkijan tehtävänä on tehdä uskottavaa tulkintaa aineistosta mahdollisimman objektiivisesti.

Alasuutari (2012) jakaa laadullisen analyysin kahteen vaiheeseen. Näistä ensimmäisellä, havaintojen pelkistämisellä, hän tarkoittaa aineiston tarkastelua teoreettisen viitekehyksen avulla ennalta päätetystä olennaisesta näkökulmasta. Toisella vaiheella, tulosten tulkinnalla, hän kuvaa, kuinka tutkimuksen kohteesta voidaan tulkinnan avulla erottaa merkitystä.

Kaikki haastattelut nauhoitettiin ja litteroitiin, jolloin niiden vertailu voitiin suorittaa laadukkaammin. Nauhoitettuja haastatteluja kerättiin yhteensä 5 tuntia 31 minuuttia. Nauhoituksista tuotettiin litteroitua aineistoa yhteensä 79 sivua. Aineistolle suoritettiin

peruslitterointi, jonka myötä nauhoitukset purettiin sellaisenaan tekstimuotoon, kuitenkin siivoten niistä puhemuotoiset ilmaisut pois. Puhemuotoisia ilmauksia haastatteluissa olivat esimerkiksi ”niinku” ja ”tota” -ilmaisut. Lisäksi turhat sanojen toistot jätettiin litteroimatta. Litterointivaiheessa myös tarkastettiin, että haastateltavan sanoma ei muuttunut puhekielisyyspoistamisen myötä.

Hirsjärvi ja muut (2007, s. 219) kuvaavat, että erityisesti kvalitatiivisessa tutkimuksessa, tutkimustulosten analysointivaihe voidaan kokea hankalaksi. Tästä johtuen he uskovat tutkijoiden tekevän analyysivalintoja jo melko varhaisessa vaiheessa tutkimusprosessia. He nimeävät aineiston ymmärtämisen menetelmäksi, jonka avulla päätelmiä tekstistä pyritään tekemään. Analyysimenetelmistä he taas mainitsevat teemoittelun olevan hyvin tyypillisesti käytetty metodi. Tässä tutkielmassa haastattelujen tuloksia pyrittiin nimenomaan ymmärtämään ja sen kautta tulkitsemaan. Teemoitteluun perustuva analyysimenetelmä valittiin käytettäväksi jo varhaisessa vaiheessa, mikä mahdollisti teema-ajatuksen käyttämisen koko tutkimusprosessissa: kysymysten laadinnassa, haastattelutilanteissa sekä analysoinnissa.

Haastattelujen litteroitu aineisto purettiin ensin osiin taulukkomalliin, jotta jokaisen haastateltavan raaka vastaus kysymykseen saatiin eriteltyä kontekstista ja sen merkitys niin sanotusti tiivistettyä. Tämän jälkeen muut vastaajan näkemykset kysymykseen liittyen eriteltiin omina huomioinaan. Näiden lisäksi kysymykset ryhmiteltiin teemoittain eri kokonaisuuksiin, joiden esittely tapahtuu luvussa 5. Teemoja analysoimalla saatiin tulkittua vastaukset ennalta määriteltyihin tutkimuskysymyksiin. Tulkinnan tulokset esitetään myöhemmin luvun 6 johtopäätöksissä.

## 5 Tutkielman tulokset

Tässä luvussa esitetään tutkielman kannalta aineiston olennaisimmat löydökset ja samalla verrataan tutkielman tuloksia aikaisempiin tutkimuksiin. Tutkimustulokset puretaan esitettyjen tutkimuskysymysten mukaisessa järjestyksessä, jotta tulosten seuraaminen on loogista. Tarkemmin avattuna, tuloksissa esitetään, mitä dataa yrityksissä käytetään ja millä sitä analysoidaan. Lisäksi kerrotaan, koetaanko kohdeyrityksissä heidän käyttämän datan ja sen analysoinnin olevan olennaista ja riittävää yrityksen päätöksenteon tueksi. Lopuksi tutkielman tuloksissa haetaan selvyyttä siihen, käytetäänkö tutkimuskohteissa data-analytiikkaa päätöksenteon tukena, mikä data-analytiikan taso on eteläpohjalaisissa yrityksissä ja miten yritykset itse näkevät tämänhetkisen tilanteensa sekä tulevaisuuden näkymät. Tuloksissa esitetään haastateltavien näkemyksiä ja kokemuksia tutkimusentekijän tulkitsemana. Suorat lainaukset on tuotu tekstiin vahvistamaan tulkintaa ja lainaukset on pyritty ottamaan mahdollisimman tasapuolisesti eri haastateltavilta. Ensimmäisessä alaluvussa esitetään taustatietoa tutkielman kohteista sekä termistöä ja viimeisessä alaluvussa esitetään taulukkomuotoiset yhteenvedot tutkielman merkittävimmistä tuloksista. Teemahaastatteluissa haastateltaville esitetyt kysymykset löytyvät tämän tutkielman liitteistä (liite 1).

### 5.1 Taustatiedot ja tutkielman lähtökohdat

Aineiston keräys toteutettiin yksilöhaastatteluilla, joita tehtiin yhteensä 6 kappaletta. Suorissa lainauksissa ja yhteenvedossa haastateltavia nimitetään lyhentein H1-H6. Kaikki yritykset toimivat eri toimialoilla. Haastateltavat työskentelevät yrityksissä, jotka harjoittavat päivittäistavarakauppaa, tukkukauppaa, vähittäiskauppaa, teollisuutta tai tuotantoa. Kaikilla haastateltavilla on pitkä kokemus alalta yleisesti, työkokemus kohdeyrityksessä vaihtelee 2–23 vuoden välillä. Haastateltavat toimivat kohdeyrityksissä johto- ja päällikkötason tehtävissä. Aihetta ei rajattu tietyn liiketoiminnan osa-alueen mukaan, joten myös haastattelut pyrittiin saamaan monipuolisesti yritysten eri liiketoimintojen parissa työskenteleviltä henkilöiltä. Haastateltavien työkuvat painottuvat yrityksen

kehittämiseen tai projekteihin, myyntiin, markkinointiin, it:hen, talouteen ja verkkokauppaan. Muutaman henkilön kohdalla rooliin sisältyy lisäksi toimintoja liiketoiminnan ja talouden tai it:n välillä.

Haastattelujen aluksi kysyttiin tutkielman pääkäsitteistä, Business intelligencestä sekä data-analytiikasta, ja tiedusteltiin mitä ajatuksia termit herättivät ennakkoon. Kunkin näkökannasta johtuen vastaukset olivat hieman hajanaisia. Pääosin aihe nähtiin kuitenkin merkittävänä perustana liiketoiminnalle.

*”No sanotaanko näin, että se on kyllä laaja aihe. Et, ja varmasti kaikki ymmärtää sen vähän eri tavoin. Riippuen, että miten on sitä ite toteuttanut tai kuinka syvällä on asiassa, että, mut se on varmasti aiheena sellanen, et se kyllä tällä hetkellä kaikkia kiinnostaa, et jos ei oo vielä oikeen omassakaan yrityksessä lähdetty mitenkään edistämään sitä, niin kyllä nyt viimestään varmasti kaikki on lähtenyt.”*  
(H2)

Kun näkökantana oli nimenomaan esitetyt termit ja niistä syntyneet ajatukset, tällöin nähtiin, että termit olivat uusia nimityksiä asioille, joita oli tehty jo kauan. Tätä tulosta tukee Shardan ja muiden (2018, s. 13) esittämä data-analytiikan termistön kehitys ja data-analytiikan elämänskaari.

*”Siinä on osittain vähän sellanen kaiku, et on keksitty hienot nimet samoille asioille mitä on aina tehty, mutta tietyllä tavalla siinä on kyllä sellanen ihan todellinen erokin, elikkä tekninen kehitys ja sanotaan tämmönen datan hallinnan helpottuminen, on tehnyt sitä, että nykyään pystytään aidosti analysoimaan asioita, jotka ei oo aikasemmin ollut mahdollisia, että siitä on tietyllä tavalla tullut tämmönen ehkä oma osansa sinne it:n ja sen operatiivisen toiminnan välimaastoon.”* (H4)

Vastaajien aito näkemys aiheesta haluttiin tuoda esiin ja jotta heidän vastauksensa olisivat mahdollisimman autenttisia, heille selvennettiin vasta haastattelun loppuksi, mikä

tämän tutkielman aineiston pohjalta oleva tutkimusnäkemys BI:stä ja data-analytiikasta on. Tämän johdosta vastauksissa tuli ilmi myös eriäviä näkemyksiä. Eräs vastaaja näkikin data-analytiikan näkökulmasta, joka tässä tutkielmassa mielletään BI:ksi. Vastaaja näkemys oli, että data-analytiikka toimii pohjana BI:lle. Samassa yhteydessä ennustava analytiikka sekä ulkoisten datalähteiden hyödyntäminen mainittiin tulevaksi trendiksi, jotka taas tässä tutkielmassa esitetään data-analytiikkana.

*”Noo, siis analytiikka joo, sehän on, siis mun mielestä raportointi ja analytiikka, jolla analysoidaan mennyttä aikaa, niin se on vähän mun mielestä jälkeen jäämässä oleva trendi, et toki me tarvitaan erilaista analytiikkaa, jolla katsotaan, että miten on mennyt ja mistä, miten liiketoiminta toimii eri osa-alueilla, mut sitten on, nythän on tulossa vahvasti semmonen trendi, että pitäis pystyä aikaisempien tapahtumien perusteella arvioimaan mitähän seuraavaksi mahtaisi tapahtua ja siihenähän pitäis yhdistää näitä muitakin tietolähteitä, kuin vain oman yrityksen hallussa olevia tietolähteitä.” (H5)*

Joissain aikaisemmissa tutkimuksissa ajatukset termien tarkoituksista ovat olleet samansuuntaiset. Davenport ja Harris (2007, s. 26) käsittelivät BI:ä isona kokonaisuutena ja kertoivat data-analytiikan sisältyvän siihen. Toisaalta lisäkysymysten myötä vastaaja kertoi, että he ovat vasta ottamassa BI:ä käyttöön ja alkuun se sisälsi heillä vain sisäistä dataa ja järjestelmissä olevaa historiatietoa. Tämä tulos osoittaa, että BI:n suhteen näkemykset sekä vastaajan, että tämän tutkimusasetelman osalta olivat yhtenevät.

Haastateltavilta kysyttiin myös, käyttävätkö he termejä Business intelligence tai data-analytiikka omassa työympäristössään. Tämänkin kysymyksen osalta vastaukset olivat eriäviä. Kahdessa yrityksessä käytettiin BI:ä, kun taas eräs vastaaja kertoi data-analytiikan olevan käytetympi termi. Vastaajista puolet eivät kuitenkaan tunnistanee kumpakaan termiä käytettäväksi työyhteisössään. He kertoivat vaihtoehtoisesti puhuvan esimerkiksi analysoinnista tai raportoinnista. Termien suhteen saatu hajanainen tulos selittyy tieteellisen tutkimuksen kautta, jossa ei tunnisteta data-analytiikkaa varsinaiseksi

termiksi. Tutkimusaineiston perusteella voidaan kuitenkin todeta, että päivittäistavara-kaupassa, termi on käytössä käytännön tasolla. Tulosten esittelyssä on kirjavan käytännön termistön sijaan käytetty yhtenäistä data-analytiikan termiä.

Data-analytiikan termi kuitenkin tunnistettiin BI:ä käyttävien tahojen keskuudessa. Näistä kahdesta vastaajasta toinen koki, että he eivät olleet analytiikan kehityksen suhteen vielä niin pitkällä, että voisivat sanoa tekevänsä data-analytiikkaa. Toinen vastaaja taas tunnisti, että hänen organisaatiossaan data-analytiikkaa tehtiin jo osaksi, käsitteenä se ei kuitenkaan ollut käytössä.

*”Data-analytiikka-termiä me ei hirveesti käytetä ja meillä ei tietysti tällaisia analytiikkoja ole sinällään, vaikka käytännössä moneen tehtävään se hyvin vahvasti kuuluu.” (H4)*

Ylipäänsä tieto päätöksenteon pohjana vaikutti olevan kaikkien haastateltavien organisaatioissa käytössä.

*”Se on vähän, ku ajais autoa ilman nopeusmittaria, jos ei oo tietoa. Ei voi olla varma, että meneekö ylinopeutta vai eikö ja onko liikenteessä tulppana.” (H5)*

## 5.2 Datan lähteet ja analysointityökalut

Ensimmäisen tutkimuskysymyksen selvittämiseksi haastateltavilta kysyttiin, mistä tiedonlähteistä he keräävät datansa ja millaisten ohjelmien avulla sitä analysoidaan. Vastauksista voitiin tehdä jako sisäisen ja ulkoisen datan välille. Sisäisistä datalähteistä mainittiin ERP ja myynninohjausjärjestelmä. Vaihtoehtoisesti sisäisenä tietolähteenä toimi tietovarasto, josta dataa haettiin jalostettavaksi edelleen. Sisäisestä datasta mainittiin asiakas-, myynti-, osto-, transaktio-, sopimus- sekä tuotantodata. Ulkoista dataa kerättiin esimerkiksi viranomaislähteistä, yhteistyökumppaneilta, tavarantoimittajilta ja tukkuliikkeiltä, internetsivustoilta, sensoreista ja julkisista tai ostetuista lähteistä, kuten

markkinatutkimuksista. Ulkoisesta datasta esimerkkeinä mainittiin erilaiset rekisterit ja kuluttajadata. Sensoridatasta esimerkkeinä mainittiin sähkön- ja vedenkulutus sekä muiden erilaisten koneiden tuottama data.

Kaikki haastateltavat käyttivät hyväkseen sisäistä dataa, ja vastaajien painopiste oli enimmäkseen sisäisen datan käytössä ja siinä, miten sitä käytettiin hyödyksi. Varsinkin myyntidata nostettiin useassa haastattelussa esille ja sen hyödyntäminen nähtiin tärkeäksi.

*"Kyllähän myyntidata on tavallaan se, minkä kautta me pohjataan se, mikä on ehkä se ydin siinä. Mitä myydään, miten myydään ja se, että sitä kautta tietysti analysoidaan hyvin montaa. Me kerätään tietysti meidän ERP:istä ja myynninohjausjärjestelmistä tietoa, jota sitten analysoidaan oikeastaan lähestulkoon kaikkien kaupallisten päätösten pohjalle." (H1)*

Ulkoisen datan käyttö voi tuoda suurempaa hyötyä päätöksenteon tueksi (Kwon ja muut, 2014, s. 392). Tästä syystä tutkielman kannalta oli tärkeää selvittää, mitä ulkoista dataa yrityksissä käytetään. Vastaajat toivat aluksi esiin lähinnä yrityksen sisäistä dataa, joten ulkoisen datan osalta haastatteluissa oli tarve esittää tarkentavia kysymyksiä. Tämän myötä ulkoisen datan käytöstä tehtiin hyviä huomioita. Lähtökohtaisesti vastaajat mainitsivat hyvin erilaisia ulkoisen datan lähteitä ja hyödyntämistarkoituksia. Tämä selittyi osaksi sillä, että yritykset toimivat eri toimialoilla. Yksi yhtenevä ja useammassa organisaatiossa käytetty tekijä oli asiakasdata. Asiakasdata voidaan sisällyttää niin sisäiseen kuin ulkoiseen dataan, riippuen kerätäänkö sitä itse ja omista sisäisistä järjestelmistä vai onko se ulkoisista lähteistä hankittua. Ulkoista asiakasdataa kerrottiin kerättävän erilaisista rekistereistä, kuin yhteistyökumppaneiden kautta. Kaikki eivät kuitenkaan asiakasdataa hyödyntäneet vaan käytössä oli pelkästään määrällistä dataa.

*"Me ei mitään profilointidataa ja tämmöstä me ei käytetä, me käytetään vaan numeraalista dataa" (H3)*

Toinen useammassa haastattelussa esiin tuotu ulkoinen lähde oli sensoridata. Sensoridataa kerrottiin hyödynnettävän jo jonkin verran, mutta useampi vastaaja näki sen hyödyntämisen kasvavan jatkossa. Sensoridataa oli tulossa käytettäväksi enemmän esimerkiksi investointien myötä ja sen hyödyntämisen kerrottiin mahdollistavan myös niihin kytkeytyneen liiketoiminnan parempaa ennustamista.

Haastateltavilta tiedusteltiin myös sisäisen ja ulkoisen datan käyttösuhdetta omassa organisaatiossaan. Kysymys oli monelle haastateltavalle yllättävä ja he kertoivat, ettei asiaa ollut heidän organisaatiossaan pohdittu. Kaikki painottivat, ettei heillä ollut tarkkaa tietoa suhteesta, eikä asia muutenkaan vaikuttanut olleen esillä yrityksissä. Vastaukset siis perustuivat heidän omiin oletuksiinsa. Tämän tiedon myötä organisaatioiden ulkoisen datan käyttöä oli kuitenkin mahdollista vertailla jossakin määrin. Optimistisin ulkoista dataa keräävä organisaatio arvioi sisäistä dataa olevan jopa puolet kaikesta käytetystä datasta, kun taas toisessa ääripäässä oli yritys, jossa vastaajan mukaan ei käytetty ulkoista dataa lainkaan. Kiinnostavaa olisi ollut myös selvittää, mikä strukturoidun, strukturoimattoman ja semistrukturoidun datan osuus oli yrityksessä, mutta tämän koettiin olevan liian hankala arvioida ilman, että tutkielman luotettavuus vaarantuisi.

Vastauksiin sisäisestä ja ulkoisesta datasta vaikutti oletettavasti haastateltavan oma datan käyttökokemus ja edustamansa osasto tai osastot, joissa dataa käytettiin. Datan käytön erot olikin havaittavissa vastauksissa.

*”No esimerkiksi markkinointiosastolla on ulkoinen data käytössä, mut sitten taas talon sisällä esimerkiksi myyntiosasto turvautuu hyvin pitkälle siihen mitä tulee ERP:istä ja mitä ERP tuottaa meille.” (H6)*

Hankaluuksia tuotti myös sisäisen ja ulkoisen datan määrän arvioiminen. Niillä nähtiin olevan isoja sisällöllisiä eroja. Vastaajat kertoivat, että ulkoinen data ei ollut samalla tavalla yhdisteltävissä tai yhtä tarkkaa ja säännöllistä, kun sitä verrattiin sisäiseen dataan.

*”Vaikkapa nyt ajatellaan joku myyntidata, niin meillä voi olla se ulkoinen myyntidata, mitä me saadaan jotain, mitä se sitten ikinä onkaan, eri lähteistä, niin se voi olla, sen tarkkuustaso sellanen, että se, et meillä on vaikkapa 5 000 riviä, niin siitä samasta 5 000.sta rivistä meillä voi olla omaa dataa 500 000 riviä, vaikka se on sama tavallaan, sama kokonaisuus, mutta meillä se tarkkuustaso on niin valtavan paljon suurempi omasta datasta.” (H4)*

Käytettävästä datasta tehtiin paljon huomioita. Ensinnäkin sisäistä ja ulkoista dataa ei nähty erillisinä tietolähteinä, vaan niitä kerrottiin yhdisteltävän ja sen kautta datan tuovan varmuutta päätöksentekoon. Ulkoisen datan saatavuuden nähtiin myös kasvaneen viimeisten vuosikymmenten aikana ja sen tehostavan ulkoisen datan käyttöä.

Big dataa ei kukaan vastaajista ottanut esille. Haastatteluissa ei tästä myöskään erikseen tiedusteltu, koska tämän tutkielman kannalta ei ole merkityksellistä, kokevatko kohdeyritykset käyttävänsä dataa vai BD:a. Käytettävää dataa kuitenkin kuvailtiin esimerkiksi valtavaksi määräksi, mikä taas tutkimusten mukaan (Chen ja muut, 2012, s. 1166; Gartner, 2016; McKinsey Global Institute, 2011) on BD:lle tunnuksenomaista. Datan ja BD:n raja on kuitenkin hyvin häilyvä, eikä tämän tutkielman perusteella voida tehdä johtopäätöksiä käytetäänkö kohdeyrityksissä Big dataa vai dataa.

Datan analysointiin oli käytössä laajalti erilaisia järjestelmiä. Haastateltavat mainitsivat käyttävänsä apunaan useita eri järjestelmiä. Haastatteluissa mainittiin erilaiset BI-ohjelmistot, Excel, ERP, CRM, myynnin analysointijärjestelmä, automaattitilausjärjestelmä, ennustamisjärjestelmä ja verkkosivujen kävijäseurantajärjestelmä. Lisäksi muutamassa yrityksessä oli käytössä omat tietokannat, joista analysointi tapahtui keskitetysti. Jossain yrityksessä oli käytössä myös analysointiin räätälöity oma työkalu. Mainittujen järjestelmien lisäksi käytössä oli useita muita pienempiä järjestelmiä, joista haettiin tukea analysoinnille tai datan visualisoinnille. Lisäksi useampi vastaaja kertoi verkkokaupan datan analysoinnissa olevan käytössä tai olevan mahdollista ottaa käyttöön useita eri työkaluja, mutta näitä ei haastattelussa tuotu esiin.

Datan analysoinnin työkaluista yleisin oli Excel, joka mainittiin lähes jokaisessa haastattelussa. Vastaajat totesivat, että Exceleitä saatettiin tehdä milloin mistäkin asiasta, usein kuitenkin, kun tarvetta oli ad hoc -tyyppiselle analysoinnille. Yhdellä vastaajalla Excel toimi datan ainoana visualisointityökaluna.

*”Siis se visualisointi tapahtuu hyvin paljon silleen, tai yleensäkin se datan analysointi, että otetaan riittävä määrä dataa ja sitten visualisoidaan ja louhitaan sitä Excelissä.” (H5)*

Yleisesti organisaatioissa oli käytössä melko laajasti erilaisia työkaluja datan analysointia varten. Tähän oli useita syitä. Ensinnä moni mainitsi esimerkiksi ERP-järjestelmänsä analysointijärjestelmänä, koska sen kautta tarkasteltiin raportteja, joiden perusteella tehtiin päätöksiä. Toisessa ääripäässä voitiin katsoa olevan ennustamisjärjestelmät, joissa oli käytössä tekoälyä. Toinen syy laajalle työkalukirjolle vaikutti johtuvan sitä, että eri työkaluja käytettiin organisaation eri osastoilla. Esimerkiksi myynnin analysointiohjelma saattoi olla enimmäkseen myyntiosaston käytössä, kun taas verkkosivun kävijäliikennettä seurasi verkkokauppaosasto. Vaikka pyrkimys monilla vaikuttikin olevan käyttää järjestelmiä keskitetysti, se ei kuitenkaan käytännössä toteutunut.

*”Meillä on kyllä toki pyrkimys, että kaikki tavallaan, raportointi, tapahtuis tietyn järjestelmän kautta, mutta kyllähän se, sitä on tosi hankala vetää sellasta linjaa, että kyllähän sitä tapahtuu sitä analysointia ihan joka järjestelmästä, et jos aattelee niitä käyttäjiä, niin kyl ne jonkinlaisia analyseja tekee sieltä ihan CRM:istä lähtien.” (H2)*

Excelin ohella toiseksi mainituin työkaluryhmä oli BI-ohjelmistot, joita käytettiin 4:n vastaajan organisaatioissa. Lisäksi yksi vastaaja kertoi heillä tehtävän kartoitusta BI-ohjelmasta ja toiselle BI-ohjelmisto oli jo tulossa. Vastaajilla oli pääosin käytössä eri ohjelmantoimittajan BI-ohjelmistoja. Vastaajista yksi ei tunnistanut heillä käytettävän BI-

ohjelmistoa lainkaan, vaikka jälkikäteen tehdyn taustatyön perusteella kyseinen työkalu markkinoi itseään sellaisena. Kuvailun perusteella kyseessä oli melko vanhanaikainen BI-ohjelmisto, jossa tiedonsiirto ei toiminut reaaliaikaisesti eikä datan visualisointi ollut käytettävissä. Tämä osoitti, että markkinoilla olevien BI-ohjelmistojen taso on vaihteleva.

*”Muodostaa kerran yössä sen edellispäivän tilanteen, et käytännössä meillä ei ai-  
van 100 %:sesti sellasta reaaliaikasta dataa oo olemassa.” (H6)*

Eräs haastateltavista toi esiin BI-ohjelman hyödyllisyyden sen esittämistavan ja joustavuuden myötä.

*”Mutta toi (BI-ohjelmisto) on tuonut sen sellasen eron siihen, että jossain johtoryh-  
mässä, kun esität lukuja, niin sä pystyt näyttämään sen samalla, että kun joku kysyy,  
et aa s-ryhmä noin, mites k-ryhmä sitten, niin sit sä pystyt vaihtamaan sen heittä-  
mällä, kun se vanha juttu on ollut se, miten se ennen on mennyt, että no mä lähe-  
tän sitten tai että no odotas, katotaas noi muut esitykset tos, mä ettin sillä aikaa  
sen.” (H4)*

Toisaalta BI-ohjelmiston joustavuus ilmeni toisessa haastattelussa huonona puolena. Vastaajan mielestä liika joustavuus ja käyttäjälle sallittu datan muokkausmahdollisuus mahdollistivat datan esittämisen virheellisestä näkökulmasta.

Tuloksien kanssa samassa linjassa on Nykäsen ja muiden (2016, s. 40) tutkimus, jonka mukaan taulukkopohjaiset työkalut ja tiedon visualisointijärjestelmät ovat yrityksissä eniten käytetyimmät työkalut. Tutkielmassa esiin nousseista visualisointijärjestelmistä erityisesti BI-ohjelmat nousevat esiin yhtenäisenä työkaluna, mutta myös lisäksi vastaa-  
jista muutama mainitsi heillä olevan käytössä myös erilliset visualisointiohjelmat.

Kaikki vastaajat eivät kokeneet BI-ohjelmistoa niin tarpeelliseksi. Ohjelmiston kartoitusta tehneen yrityksen haastateltava ei uskonut järjestelmän tuovan vastauksia heidän

haasteisiinsa. Hän pohti erityisesti myynnillisiä seikkoja, kuten sitä, auttoiko järjestelmän tuottama tieto tekemään lisää myyntiä tai oivaltamaan asiakkaasta jotain sellaista, joka auttaisi myynnin edistämisessä.

Toinen, vasta BI-ohjelmiston käyttöönottoa valmisteleva haastateltava taas kuvaili ohjelmiston avulla olevan helpompi hallita tietoa. Vastaajalla oli haastatteluhetkellä käytössä Excel päätoimisena analysointityökaluna ja sen haittapuolena hän näki tiedostorakenteen sekä käyttäjälle sallitun vapaan tietojen muokkauksen, joka aiheutti virheellisyyttä. BI-ohjelmiston myötä hän uskoi tiedon olevan puhtaampaa ja laadukkaampaa.

*”Sitten jos se olis suoraan joku tämmönen raportointinäkyvä, joka näyttää tiedot visualisoituna juuri semmosena, kun ne on, niin sehän olis paras ratkaisu. Silloin se olis tavallaan kiistämätöntä se tilanne, että Excelissä voi tulla virheitä, siis se käy helposti, kun isoja datamassoja käsitellään, niin voi tulla näppäilyvirheitä ja sä voit vahingossa poistaa rivejä tai lisätä jotakin mitä sä et halunnukaan, niin sen, tavallaan, et jos ajatellaan tiedon eheyttä ja kiistämättömyyttä, niin se vois olla Excelissä vähän kyseenalainen, että onko näin vai eikö.” (H5)*

Pääosin kohdeorganisaatioissa visualisoitiin dataa BI-ohjelmistojen kautta tai käytössä oli lisäksi erillisiä visualisointiohjelmia. Esimerkiksi Begum ja George (2017, s. 1651) kertoivat datan visualisoinnin olevan helppoa ja tehokasta BI-ohjelman avulla, joten selkeästi tulokset ovat yhtenevät heidän näkemyksensä kanssa. Yhdellä vastaajista datan visualisointia ei ollut lainkaan käytössä ja toisella visualisointi tapahtui ainoastaan Excelin kautta.

Tutkielman teoriaosuudessa tuotiin esiin data-analytiikan työkalut BI-ohjelmistoja mutkikkaampina työkaluina. Tutkielman tuloksien perusteella voimme arvioida, että data-analytiikan työkaluja oli käytössä lähinnä vain yrityksissä, jotka mainitsivat hyödyntävänsä ennustamisjärjestelmiä.

Yleisesti datan saatavuus ja järjestelmien kehitys oli ollut huimaa. Yksi vastaajista kertoi tilanteesta reilu 10 vuotta sitten ja kuvaili, kuinka kehitys oli muuttanut työn sisältöä. Vain paperilla oleva tieto oli lähes kadonnut, kymmenien Excelien täyttämisestä ja yhdistelemisestä oli päästy tiedon saamiseen lähes nappia painamalla ja tiedon tarkkuus oli muuttunut hyvin yksityiskohtaiseksi.

### **5.3 Datan ja analysoinnin olennaisuus ja riittävyys sekä haasteet**

#### **5.3.1 Datan ja analysoinnin olennaisuus ja riittävyys**

Kun vastaajilta oli saatu tieto heidän käyttämästään datasta sekä sen analysointityökaluista, voitiin siirtyä seuraavan tutkimuskysymyksen pariin. Sen selvittämiseksi haastateltavilta kysyttiin, miten olennaiseksi ja riittäväksi he näkivät datan sekä datan analysoinnin. Datan analysoinnilla tässä yhteydessä tarkoitetaan erilaisista järjestelmistä saatuja, niin visuaalisia, kuin numeromuotoisiaakin, raportteja, jotka toimivat datan analysoinnin välineinä.

Ensiksi haastatteluissa keskityttiin datan ja sen analysoinnin olennaisuuteen. Vastaukset olivat melko yhteneviä kaikkien haastateltavien osalta ja kaikki kokivat datan ja sen analysoinnin olevan yritykselleen olennaista liiketoiminnan päätöksenteon tueksi.

*”Kyllä. Joo ne (data ja datan analysointi) on välttämättömiä ja liiketoiminnan ohjaamiseksi.” (H5)*

Kysymystä pilkottiin pienempiin osiin, jotta vastaajat saivat vastata erikseen niin datan olennaisuuteen kuin myös datan analysoinnin olennaisuuteen. Datan olennaisuuden osalta kaikki kokivat, että data oli olennaista ja tärkeää päätöksenteon kannalta. Tämä tulos osoittaa, että data nähdään yrityksissä tärkeänä päätöksentekovälineenä. Teoreettisella puolella tuloksia tukevat tutkimukset, joiden mukaan datasuuntautunut

päätöksenteko on alettu nostamaan mukaan yritysten strategiaan (Davenport ja muut, 2001; Viitala & Jylhä 2019, s. 136). Eräs vastaaja kommentoi, että datan merkitys päätöksenteossa oli heillä isossa roolissa ja sen käyttö erityisesti viime vuosina oli ollut kasvusuuntainen.

Myös datan analysointi eli raportointi todettiin vastaajien keskuudessa olennaiseksi. Raportoinnilla kerrottiin olevan iso merkitys päätöksenteossa. Kaikilla raporteilla uskottiin olevan oma käyttäjäkuntansa ja eräs vastaaja esitti, että jokainen raportti oli aikanaan muodostettu jonkun päätöksentekijän tarpeesta johtuen.

*”Kyl mä sanoisin, että varmasti kaikki (raportointi) mitä nyt on, niin kyllähän se olennaista on. Ei oo hirveesti turhaa tehty.” (H2)*

Toisaalta datan ja siitä muodostettavan raportoinnin tarpeessa oli havaittavissa eroja. Erilaisen datan tarpeellisuudesta ei voitu tehdä jakoa, koska oli hyvin käyttäjästä riippuvaista, mikä data kunkin päätöksenteossa oli tärkeää. Haastateltava kertoi myös, kuinka jokin raportti saattoi olla käytössä päivittäin, ja toinen vain kerran vuodessa, mutta niiden välillä ei silti voitu tehdä erottelua, kumpi raporteista oli tärkeämpi.

*”Et tavallaan se tarpeellisuus ja tärkeys, niin se ei oo yhtä kuin sen käyttöuseus.” (H4)*

Myös ajoitus koettiin merkitykselliseksi elementiksi, kun puhuttiin joko datan hankinnasta tai raportin käyttämisestä.

*”Näitä tulee aina välillä, et on sellanen kuvitelma, et tää data tai tää raportti on hirveen tärkeä ja se on pakko saada, niin kyl välillä käy niin, että aika ajaa sen ohi, ennen kun se on ees saatu.” (H4)*

Vastauksista sai käsityksen, että vastaajat arvostivat sitä dataa ja datan analysointia, mitä he työssään hyödynsivät. Siksi heiltä tiedusteltiin myös, kokivatko he, että tarjolla olisi turhaa dataa tai turhaa raportointia. Datan osalta eräs vastaaja koki, että järjestelmissä on mahdollisesti jotain sellaista dataa säilössä, jota päätöksenteossa ei hyödynnetä. Toisaalta tämän myötä voidaan myös tulkita, että dataa säilytetään jonkin muun syyn takia, kuin että sitä hyödynnettäisiin päätöksenteossa, esimerkiksi lainsäädännöllisistä syistä. Toinen vastaaja totesi, että järjestelmistä ei löydy juuri hyödyntämätöntä dataa. Tämä taas viittaa siihen, että heillä järjestelmiä mahdollisesti hallitaan ja putsataan tarkemmallalla tasolla. Datan analysoinnin osalta osa vastaajista tunnisti, että aikaisemmin käytössä oli ollut raportteja, joiden tarpeellisuutta oli kyseenalaistettu. Yritykset vaikuttivat kuitenkin kaikki olevan jo siinä pisteessä, että prosesseja oli viilattu ja turhaa manuaalista koontityötä pyritti poistamaan, jolloin myös tarpeetonta raportointia oli saatu karsittua.

*”Aika paljon meillä jossain vaiheessa joku teki koontiexceleitä erilaisista asioista ja sitten ollaan kyllä pyritty hyvin kriittisesti sitä käymään läpi, että kuka tätä käyttää, mihin tarkoitukseen, ja onko tätä järkeä tuottaa tätä kyseistä dataa vai ei.” (H1)*

Kolmas vastaaja näki turhaa raportointia nousevan esiin lähinnä konserniraportoinnin osalta. Varsinkin eri maissa toimivat konserniyhtiöt saattoivat tulkita asioita hyvin eri lailla, johtuen näkökulmasta, käytännöistä tai kulttuurista. Niinpä osa raportoinnista vaikutti vastaajan osalta tarpeettomalta. Tämän vastauksen perusteella ei voitu kuitenkaan tehdä johtopäätöksiä, koska tiedossa ei ollut kuinka tärkeäksi joku raportointi toisessa konserniyhtiössä koettiin.

Haastatteluissa otettiin aiheeksi myös toinen ääripää, eli se, kokivatko vastaajat, että raportointia tulisi joiltakin osin olla enemmän. Vastaukset tähän kysymykseen olivat melko eriäviä. Eräs vastaaja koki, että raportointia oli jo liikaakin, ja laaja raportointi saattoi hankaloittaa jonkin olennaisen tiedon löytämistä. Toinen vastaaja taas ilmaisi, että raportointia tulisi olla enemmän. Vastauksista ilmeni, että haastateltava käytti raportteja päätöksentekonsa tukena, mutta ei itse laatinut niitä. Vastaaja koki, että hänen

tarkastelemansa raportit eivät aina olleet riittävän kattavia. Tämän osalta raportointia kuitenkin kehitettiin jatkuvasti raportin laatijan ja päätöksentekijän yhteistyöllä. Kolmas vastaaja otti näkökannaksi nimenomaan sen, tulisiko dataa olla enemmän ja sen osalta hän näki hyödyntämätöntä liiketoimintaa esimerkiksi sensoridatapuolella.

Seuraavaksi kysymyksissä siirryttiin datan sekä datan analysoinnin riittävyteen. Näkemykset riittävyden osalta olivat mielenkiintoisia. Pääosin vastauksista oli havaittavissa, että dataa sekä raportointia kehitettiin jatkuvasti.

*”No täähän on just semmonen, että tää on niinkun karkkikaupassa, et koska se nyt on riittävä se pussi sitten.” (H3)*

Vastaajat totesivat melko yksimielisesti, että nykytilanne oli kohtuullisen riittävä, mutta samaan yhteyteen mainittiin, että jatkuvasti asioita vietiin eteenpäin ja kehitettiin, sekä tarkennettiin, minkä myötä ilmeni uusia datantarpeita tai raportointivaatimuksia.

*”Kyllä me koko ajan sitä tarkennetaan sitä analyysia siitä, että miten isot kalat on niin sanotusti otettu, ja nyt jatkuvasti mennään pienempiin osakokonaisuuksiin siellä.” (H1)*

Ainoastaan yksi vastaajista koki, että heidän tämänhetkinen raportointinsa ei ollut riittävä. Tämä johtui tiedon yhdisteltävyysongelmista ja siitä, että vastaaja koki tiedon olevan hajallaan eri osastoilla, sen sijaan, että se olisi ollut keskitetysti päätöksentekijän käytävissä. Davenportin (2006, s. 102) tutkimuksen perusteella BI:ä hallinnoidaan osastotaisesti, joka voi aiheuttaa kuvatun laisia ongelmia ja päällekkäistä työtä. Tämän tuloksen osalta on huomioitavaa myös, että kohdeyrityksessä oli käytössä vanhanmallinen BI-ohjelmisto, joka osaltaan saattaa selittää puutteita. Myös vastaaja itse kuvasi BI-ohjelmistoon kankeaksi.

Toisaalta haastateltavien joukossa oli ainoastaan yksi yritys, jolla ei ollut käytössä, eikä tulossa BI-ohjelmistoa. Vastaja kuitenkin koki heidän nykyisen raportointitasonsa riittäväksi ja ilmaisi, ettei heillä ollut totuttu sen tyyppiseen johtamiseen, jota BI-ohjelmiston käyttäminen vaatii. Haastattelussa ilmeni myös, että vastaajan mielestä kohdeyrityksen liiketoiminta oli toistaiseksi tuottavaa ilman kehittyneempää raportointiakin.

*”Mä tiedän, että niitä on ja olisi tapoja, mutta toistaiseksi me pärjätään vielä tällä tavalla, et sanottu, niin se taas se turha numeronmurskaaminen tai tämmönen analysointikin, niin me nähdään, et me ollaan riittävällä tasolla siinä, et meidän ei tarvi mennä syvälle.” (H3)*

Oman liiketoiminnan parantaminen oli selkein havaittavain syy, miksi suurin osa vastaajista kehittivät dataa tai raportointia. Toisaalta useat haastateltavat tunnistivat kasvavan kilpailun ja kertoivat, että ilman kehitystä liiketoiminnalla oli vaara jäädä kilpailijoiden jalkoihin.

*”Joo siis jos ajatellaan yleisesti ottaen, niin siis olisi vaarallista tuudittautua tämänöseen hyvinolontilaan, että kaikki on hyvin eikä mitään tarvitse tehdä, koska muut kuitenkin painaa meneen kokoajan täysillä.” (H5)*

Tätä tulosta tukee Davenportin (2006) tutkimus, jonka mukaan yritysten kilpailukyky on kiristynyt ja keinot erottautua ovat vähissä. Hän tutkimuksessaan nimeää data-analytiikan ja sen avulla optimoidut liiketoimintaprosessit yhdeksi erottautumiskeinoksi.

Yksi vastaajista totesi, että heidän yrityksensä ei vielä muodostanut dataa kaikista toiminnoistaan ja mainitsi erityisiä toimintoja, joiden seuraamisesta datamuodossa olisi hyötyä hänen liiketoiminnassaan. Tämä osoittaa, että yrityksillä on myös tiedossa jo selkeitä kehittämiskohtia omassa liiketoiminnassaan, joissa dataa voidaan hyödyntää paremmin päätöksenteon tueksi. Toisaalta monen osalla esiin nousivat resurssit. Vaihtoehtoisia kehittämiskohteita oli paljon ja hankaluuksia oli arvioida, mikä niistä olisi

liiketoiminnan kannalta kannattavinta. Ilman selkeätä esitystä kehittämiskohteen vaikutuksista liiketoimintaan ja kannattavuuteen, ne jäivät usein toteuttamatta. Lönnqvist ja Pirttimäki (2006, s. 38) esittivätkin liiketoiminnassa lähes minkä vain olevan mitattavissa oleva tieto, mutta haasteita nousevan esiin juuri siinä, mitä yrityksen on tehokasta ja hyötypanossuhteeltaan järkevää mitata. Myös Khade (2016, s. 986) sekä Markkula ja Sävaniemi (2015, s. 37) esittivät olennaisen datan löytämisen olevan suuri haaste organisaatioille. Tutkielman tulos osoittaa myös, kuinka henkilöstöllä on suuri vaikutus kehittämiskohteiden arvioinnissa ja esittämisessä. Aikaisemmissa tutkimuksissa Maisel ja Cookins (2014, s. 69) tunnistavat henkilöstön roolin sekä yrityskulttuurin yrityksen menestekijöiksi. Heidän tutkimuksessaan muutoshallinnalla sekä käyttäytymismuutoksella on suuri vaikutus menestyksen tavoittelussa.

Muutamit vastaajat pohtivat omaa asemaansa suhteessa kilpailijoihin. Vastauksissa korostui, että he uskoivat olevansa data-analytiikassa jäljessä kilpailijoihin verrattuna.

*”Tää analysointi onkin loputon suo ja tiedän, että meitä edellä tässä on paljonkin pelureita” (H3)*

Eräs vastaaja kertoi isojen kilpailijoiden data-analytiikan hyödyntämisen olevan tehokkaampaa, koska sama data-analytiikka oli monistettavissa esimerkiksi 10-kertaisesti suuremmalle yksikkömäärälle, kuin mille se vastaajan organisaatiossa tehtiin. Niinpä pienemmässä kokoluokassa data-analytiikan hyödyntäminen ei ollut yhtä kannattavaa ja useassa yksittäisessä data-analytiikan hyödyntämiskohteessa pohdittiin sen tuomaa hyötyä suhteessa hintaan.

*”Tässä tulee justiin se resurssikysymys siitä kannattavuudesta, että se ois varmaan kannattavampaa, tai siis tietenkin olis kannattavampaa, jos niitä yksiköitä olis enempi, me voitais mennä tarkemmalle, tällä yksikkömäärällä meidän pitää vetää raja johonkin.” (H1)*

Tämä havainto kiteyttää hyvin sen, miksi tutkielman tulos datan ja sen analysoinnin riittävyydelle on esitetyn lainen. Yrityksissä niiden taso on riittävä, mutta silti samalla sen ei koeta milloinkaan olevan riittävä. Tämän voi tulkita johtuvan siitä, että yrityksissä on vedetty näkymätön raja datan ja sen analysoinnin riittävyydelle omassa ympäristössä, jolloin tietyn tasoista materiaalia voidaan pitää riittävänä.

Yritykset vaikuttivat olevan toisiinsa nähden hyvin eri vaiheissa data-analytiikan hyödyntämistä. Vaikka raportoinnin riittävyden osalta suurin osa vaikutti tyytyväisiltä tilanteeseen, yksi vastaaja mainitsi ongelmista oman sisäisen datan haltuunotossa. Vastaaja kertoi esimerkkinä, että raportit olivat melko pintapuolisia ja niiden perusteella ei voinut suorittaa syvempää tarkastelua.

*”Mut se mikä meillä on hyvin puutteellista itsellä tällä hetkellä, on täs raportoinnissa se, että sanotaanko niin, että meillä on dataa ja meillä on historiadataa hyvin hyvin pitkälle taaksepäin, mut mun mielestä meillä ei osata hyödyntää sitä tällä hetkellä.” (H6)*

Tutkielman teoriaosuudessa käsiteltiin datan ja analysoinnin laatua ja niitä tiedusteltiin myös haastateltavilta. Sisäisen datan laatua pidettiin kohdeyrityksissä pääsääntöisesti hyvänä tai erinomaisena, ulkoisen datan laatua kuvailtiin hyväksi. Jos sisäisen datan osalta järjestelmissä olikin huonolaatuista dataa, tämä vaikutti olevan vastaajien tiedossa, joten näin he joko tekivät päätöksiä datan puutteet tiedostaen tai olivat hyödyntämättä kyseistä dataa. Laihonen ja muut (2013, s. 19–20) kirjoittivat sisäisen datan olevan usein heikkolaatuista ja varsinkin ydindatan vaativan puhdistamista ja laatutarkastusta. Tässä tutkielmassa ei ollut havaittavissa ongelmia sisäisen datan laadun suhteen. Vastauksista kyllä tuli esiin, että sisäistä dataa hallittiin ja esimerkiksi auditoitiin, joten tämän perusteella voidaan olettaa, että sisäisen datan mahdolliseen huonoon laatuun on jo aiemmin puututtu sen ilmentyessä.

*”Pitkän harjoittelun jälkeen, niin se (data) on luotettavaa.” (H5)*

Ulkoisen datan osalta sen laatua pidettiin myös hyvänä. Kwonin ja muiden (2014, s. 391) tutkimuksessa datan laatuun ei vaikuttanut se, oliko se sisäisestä vai ulkoisesta datalähteestä kerättyä. Tämänkään tutkielman tuloksissa laatuero sisäisen ja ulkoisen datan välillä ei vaikuttanut olevan merkittävä, kuitenkin sillä erotuksella, että sisäistä dataa kutsuttiin joidenkin haastateltavien keskuudessa myös erinomaiseksi.

Pohdintaa laadusta alkoi tulla enemmän, kun keskustelu siirtyi raportoinnin laatuun. Vaikka data itsessään koettiin hyvälaatuiseksi, useampi vastaaja mainitsi, että ongelmia aiheutti datan kääntäminen raportointimuotoon ja sen myötä raportin esittäjän tulkinta ja visualisointi datan tuottamasta tiedosta. Tämä saattoi ilmetä joko huonolaatuisena tai jopa virheellisenä raportointina.

*”Ehkä se tulokulma ja se datan tulkinta riippuu siitä, et joo, et kuka sen esittää ja millä tulokulmalla sitä tuodaan, ja se tietysti vaatii vähän seurantaa.” (H1)*

Eräs vastaaja koki, että heidän käyttämänsä BI-ohjelmisto mahdollisti esimerkiksi osamattoman käyttäjän käsissä, väärän tiedon levittämisen organisaatiossa. Virheellisen tiedon leviämistä olikin pyritty hallitsemaan ja eliminoimaan yhtenevillä toimintatavoilla ja muokkausoikeuksien rajaamisella.

Eräs vastaaja kertoi, että ulkoisen datan sisältö aiheutti enemmän epävarmuutta. Sisäisen datan osalta oli tiedossa hyvinkin tarkkaan mistä se oli muodostunut. Toinen vastaaja kommentoi datan laadun olevan riippuvaista siitä mistä tieto muodostui. Vastaajan näkemys oli, että suurin osa datasta muodostuu ihmisten tekemistä toiminnoista ja ihmisluonteen mukaisesti, ihmisen tuottama data sisältää inhimillisiä virheitä.

Kun datan ja datan analysoinnin olennaisuuteen, riittävyteen, tarpeellisuuteen ja laatuun oli saatu riittävästi näkökulmia, haluttiin vielä tuoda esiin motiivit, minkä perusteella dataa ylipäänsä kerättiin ja analysointiin. Davenport (2013, s. 122) jakaa data-

analytiikkapohjaisen päätöksenteon vaiheisiin, joka lähtee liikkeelle ongelman tunnistamisesta tai kysymyksen esittämisestä, johon datan avulla haetaan vastausta. Haastateltavalta tiedusteltiin, miten heillä oli kartoitettu päätöksentekoprosessi, oliko ensin pohdittu mitä dataa tarvitaan päätöksen tueksi ja vasta sitten lähdetty jalostamaan sitä vai oliko päätöksiä lähdetty tekemään saatavilla olevan datan perusteella. Eräs vastaajista kertoi, että päätöksentekoprosessi heillä oli vasta lähenemässä kuvattua laista prosessia. Tähän asti tieto oli muodostunut paloittain, eikä kaikkea dataa aina muodostettu päätöksentekijän tiedontarpeesta johtuen.

*”Et kyl varmasti on pyritty siihen, mut ei oo kyl vielä oikeesti lopullisesti toteutettu, et se ois käyty sen prosessin kautta läpi.” (H2)*

Kukaan vastaajista ei maininnut, että päätöksentekoprosessi heillä perustuisi strategiaan. Toisen vastaajan kuvailu kuitenkin sopi yhteen päätöksenteon mallin kanssa. Vastaaja kertoi, että heillä tietotarpeet tulevan suoraan päätöksentekijöiltä, jolloin vaikuttaisi, että tämän vastaajan osalta tulos tukee Davenportin (2013, s. 122) tutkimusta. Lisäkysymysten myötä kuitenkin selvisi, että prosessi toteutui vain sisäisen datan hyödyntämisessä. Ulkoisen datan osalta haastateltava kuvaili prosessin olevan hajanaisempaa. Taivotteena ulkoisen datankin osalta vaikutti kuitenkin olevan prosessin aloittaminen päätöksentekijän tarpeista. Muutamat vastaajista toivat esiin datan kerääjien sekä päätöksentekijöiden välisen suhteen. Vastaajilla oli organisaatiossaan oma it-osasto, joka hallinnoi sekä tuotti dataa päätöksentekijöille. Ensimmäinen haastateltavista kertoi, että näkemykset it-osaston ja päätöksentekijän tarpeiden välillä eivät aina kohdanneet.

*”Ehkä näin yleisesti sanotaanko, että it-osastolla tai it-ihmisillä saattaa olla vähän erilaiset intressit, tai varsinkin, jos siellä ei oo vahvaa kaupallista näkemystä tai semmosta, niin sieltä saattaa sitten myös tulla näkemyksiä, jotka nimenomaan ei sitä kaupallista välttämättä niin paljon, tai jolla ei kaupallisella puolella nähdä yhtä suurta merkitystä, kun ehkä sitten sillä it-puolella.” (H1)*

Tässä organisaatiossa vastuu datan hallinnasta oli aikaisemmin ratkaistu niin, että it-osasto oli itsenäisesti tuottanut dataa, jota oli tarjottu päätöksentekijän käyttöön. Myös uudet projektit olivat olleet it:n hallinnoimia, jolloin ne eivät aina palvelleet suoraan liiketoimintaa. Tämä toimintatapa olikin todettu epäedulliseksi. Vastaaja kertoi tänä päivänä prosessin toimivan niin, että kaupallinen puoli ohjaa it-osastoa, jolloin tarve datalle tai raportoinnille tuli aina päätöksentekijän taholta.

Toisen haastateltavan rooliin kuului toimia kytköksenä it:n ja liiketoiminnan välissä. Tämä vastaaja kertoi, että informaatiota tuli molemmista suunnista. Päätöksentekijältä saattoi tulla tietotarvepyyntö, kun taas it-osastolta saattoi tulla tietoa mahdollisuuksista pääätä käsiksi uuteen dataan. Eri tahojen tarpeet eivät aina kohdanneet, joten sen johdosta tässä välissä toimi tulkkeja, jotka huomioivat päätöksentekijän tarpeet ja it-osaston kapasiteetin ja yhdistivät molempien mahdollisuudet parhaaksi ratkaisuksi.

Muutama vastaaja kertoi, että data-analytiikka nähtiin koko yrityksen vaikuttamiskeinona ja uusien ideoiden syntymistä ei haluttu rajoittaa tietyn osaston tai päätöksentekijän työkaluksi. Eräs vastaaja ilmaisi heillä olevan matala organisaatio, jossa asioita kyseenalaistettiin eri tasoilta. Tämän myötä hän kertoi saadun paljon kehitystä aikaan.

Puhuttaessa dataperusteista liiketoiminnasta ja päätöksenteosta, haastatteluissa nousi esiin myös sen vastakohta, päätöksenteko mututuntumalla. Erityisesti vastauksissa korostui, että päätöksiä oli aikaisemmin tehty mututuntumalla, mutta nykytilanne perustui enemmän dataperusteiseen päätöksentekoon. Muutama vastaaja kuitenkin myönsi, että tilanteita tuli edelleen päivittäin, jolloin dataa ei ollut käytettävissä. Tällöin päätöksenteossa turvauduttiin mututuntumaan, mutta päätöksenteko perustui yleensä myös vankaan ammattitaitoon ja henkilöiden havainnointiin organisaationsa toiminnasta. Eräs vastaajista kuvasi heillä tehtäviä päätöksiä hybridimalliksi, jonka mukaan mututuntumalla tehdyt päätökset ja inhimilliset virheet kuuluivat osaksi päätöksentekoa.

*”Tämäkin on aika uutta meillä, että lukuja analysoidaan näinkin tarkasti, verrattuna mihinkä on totuttu, että, mutta se ala uudistuu ja sitä kautta otetaan enemmän ja enemmän dataa tueksi päätöksenteolle.” (H3)*

Yksi Kironin ja muiden (2012, s. 48) tutkimuksessa ilmennyt datan käyttöä päätöksenteossa rajoittanut syy oli, että tarvittava data ei ole aina saatavilla. Heidän tutkimuksensa vahvistaa ilmenneitä tuloksia. Vastaajien tahtotila on kyllä perustaa päätöksensä dataan, mutta sen ei koeta olevan aina mahdollista.

Haastatteluissa kävi ilmi, missä osastoissa tai mitkä henkilöryhmät käyttivät eniten data-analytiikkaa. Joka haastateltavan organisaatiossa sitä nähtiin käytettävän eri osastoilla eli vastauksissa oli keskenään paljon eroja. Ainoastaan talousosasto oli vastaus, joka toistui useammassa haastattelussa. Yhdessä organisaatiossa käyttö painottui johtoasemassa toimiviin henkilöihin sekä talousosastoon, toisessa käyttö oli suurinta myynnissä ja markkinoinnissa, kolmannessa kerrottiin, että myynti, talous sekä tuotanto käyttivät eniten data-analytiikkaa ja neljännessä suurimmat hyödyntäjät olivat talous ja materiaalihalinto. Schmarzo ja Partlow (2013, s. 4–5) nimesivät markkinointiosaston yhdeksi data-analytiikan käytöstä hyötyväksi tahoksi. Ennako-oletuksena tutkielmassa oli, että markkinointiosasto olisi myös noussut tuloksissa esiin useammin. Tämä kuitenkin saattaa selittyä osaksi sillä, että kaikkien henkilöiden toimenkuva ei yltänyt markkinointipuoleen ja heillä ei välttämättä ollut kaikkea tietoa siellä käytettävistä menetelmistä.

Selkein selitys tälle on, että organisaatiot käyttävät data-analytiikkaa paikoissa, joissa siitä todetaan olevan eniten hyötyä. Myös vastaajan oma asema yrityksessä ja sen myötä näkökulma on saattanut vaikuttaa vastauksiin. Koska organisaatiot toimivat hyvin erilaisilla toimialoilla, myös liiketoiminnan haasteellisimmat paikat olivat toisistaan paljolti poikkeavia. Schmarzo ja Partlow (2013, s. 4–5) nimesivät data-analytiikan käytöstä eniten hyötyviä osastoja. Niistä esimerkiksi hankinta, tuotekehitys, hinnoittelun ja henkilöstöhallinto olivat osa-alueita, jotka eivät tämän tutkielman haastatteluissa nousseet esiin.

Vastaajat toivat esiin nimenomaan osastoja, joissa uskoivat data-analytiikkaa hyödynnettävän eniten. Toisaalta vastauksien perusteella oli myös selkeää, että sitä ei hyödynnetty joka osastolla läpi yrityksen. Samankaltainen tulos ilmeni Nykäsen ja muiden (2016, s. 33) tutkimuksesta, jossa vain joka kuudennen vastaajan mukaan BI:ä käytettiin yrityksissä joka osastolla.

### 5.3.2 Liiketoiminnan ja data-analytiikan haasteet

Haastattelun aikana vastaajia pyydettiin kuvailemaan erilaisia haasteita, joihin he olivat törmänneet oman organisaationsa liiketoiminnassa. Samalla tiedusteltiin myös data-analytiikan käytössä ilmenneistä haasteista. Tämä kartoitus tehtiin, jotta saatiin parempi tuntemus kohdeyrityksistä sekä heidän liiketoiminnalleen tyypillisistä ongelmakohtista. Lisäksi tämä auttoi luomaan paremman kokonaiskuvan yhtiön data-analytiikan tilasta.

Eräs vastaajista kuvasi melko kokonaisvaltaisen liiketoiminnan haasteen. Analysointijärjestelmien käyttöönotto ja sen myötä data-analytiikan käyttöönotto oli organisaatiossa ollut merkittävä muutos. Vastaaja kuvasi, kuinka muutos oli parantanut heidän liiketoimintaansa datasuuntautuneeksi ja tuonut mukanaan valtavia hyötyjä, kun arvailujen sijaan päätöksiä voitiin perustaa datalle.

*”Semmonen merkittävä vaikutus siitä, mistä mututuntumasta tiedolla johtamiseen tai datan analysointiin siirryttiin. Ja se on nyt sitten tietysti jatkuu ja tuo koko ajan tarkempaa ja tarkempaa tietoa siihen tekemiseen.” (H1)*

Toinen vastaaja nimesi yhdeksi heidän yleiseksi haasteeksensa sen, että tarvittava data ei aina ollut saatavilla. Tämä tulos oli mielenkiintoinen, koska vastaaja kuitenkin koko, että yleisesti heillä datan riittävyys oli hyvällä tasolla. Voidaankin tulkita, että datan puuttuminen tuli ilmi yllättävissä tai harvemmin tapahtuvissa päätöksissä. Toinen mahdollisuus on, että yrityksen kehittäessä toimintaansa ja datavarastoaan, se on törmännyt datan saatavuusongelmiin.

Eräs vastaaja kertoi kilpailijoiden olevan yksi heidän suurimmista haasteistaan. Kilpailijoiden seuranta oli tärkeää heidän liiketoiminnalleen ja vastaaja kokikin, että tähän oli saatavissa apua tietokannoista, joiden avulla voitiin analysoida kilpailijoiden toimintaa.

*”Tullaan siihen, että pitää olla paljon hektisempää ja pitää olla oikeaa tavaraa oikeeseen aikaan, oikeessa paikassa. Sen takia pitää pystyä analysoimaan markkinaa kokonaisuudessaan, että kenellä on, mitä ja missä ja mihinkä hintaan ne myyvät niitä, ettei ole enää se oma pesä, vaan katsotaan enemmän ulos myös.” (H3)*

Myös toinen vastaaja nosti keskustelussa esiin kilpailijat ja kertoi, että kilpailijoiden toimintaa kyllä seurattiin, mutta heillä toiminta ei ollut niin säännöllistä, että seurantaa olisi tehty data-analytiikan avulla.

Kokonaisvaltaisten haasteiden sijaan moni vastaajista toi esiin pienempiä haasteita ja ongelmakohtia liiketoiminnassaan, joissa data-analytiikkaa hyödynnettiin. Analysointi nähtiin jatkuvana työnä ja päivittäisen päätöksenteon osana. Eräs vastaaja kertoi markkinoinnin ja siihen panostamisen olevan jatkuva haaste, johon myös saatiin apua data-analytiikasta. Toisaalta järjestelmät, joista tietoa haettiin, alkoivat vanhentumaan, joten tämä taas aiheutti organisaatiolle uusia haasteita. Haastateltava kertoikin joidenkin kilpailijoiden vastaavan järjestelmien vanhenemiseen kehittämällä uusia omia järjestelmiä.

Tuotantoa harjoittavan yrityksen puolella haasteita ilmeni tuotannon prosessien maksimoimisessa. Data-analytiikan avulla nähtiin mahdollisuuksia saavuttaa tehokkuutta läpi linjan.

Yksi vastaaja kertoi heidän haasteenansa olleen asiakastuntemus ja esimerkki kertoikin, kuinka heillä oli data-analytiikan avulla opittu tuntemaan asiakas paremmin. Heillä asiakkaan tuntemiseen vaikutti vahvasti asiakkaan ympäristö ja kulttuuri, joten vahvan

tietopohjan avulla he saavuttivat parempaa tuntemusta omista asiakkaistaan ja sen kautta pystyivät hyödyntämään tietoa liiketoimintapäätöksissä.

*”Näiden asioiden analysoimiseksi ja päätelmien tekemiseksi, niin näistähän päästään raportointijärjestelmien avulla ja analytiikkajärjestelmien ja analytiikalla kiinni, että mikä oikeesti kannattaa ja mihinkä leikkiin kannattaa iskeä ittensä mukaan.” (H5)*

Haastateltavilta kysyttiin myös heidän kohtaamistaan haasteista, joita he olivat kokeneet data-analytiikan käytössä. Yksi vastaajista kertoi haasteita ilmenevän henkilöstöresursseissa. Eritoten sellaiset henkilöt, joilla oli ymmärrys sekä liiketoiminnasta, että data-analytiikasta, olivat vähissä. Vastaaja kertoi, että tähän asiaan ei ollut vielä löydetty ratkaisua.

*”No kyl meillä varmaan suurin haaste on ihan tämmöses osaamises ja ymmärryses. Et se on hyvin pieni porukka, joka ymmärtää sen, että mitä täältä vois saada irti ja osaa sen suhteuttaa sit siihen, et mitä oikeestaan tarvii.” (H2)*

Toinen vastaaja nimesi haasteita myös riittämättömissä resursseissa. Haastattelussa ilmeni, ettei resurssien puute kuitenkaan koskenut henkilöstön saatavuutta vaan enemmänkin sitä, että data-analytiikkaan ei haluttu yrityksessä investoida.

Useampi vastaaja kuvaili henkilöstön roolia data-analytiikassa tärkeäksi. Kuvailuissa korostui, että data-analytiikka oli oiva päätöksenteon apuväline, mutta ihmisen osuus päätöksenteossa ei ollut korvattavissa pelkästään data-analytiikalla.

*”Kuitenkin se pitää muistaa, että ihminen tekee ne päätökset aina, että tää (data-analytiikka) on vain avuksi, et siinä vaiheessa, kun se tekoäly rupee päättämään sen, et laitetaanks toi tehdas kiinni vai ei, niin sit mä kyllä, sit ollaan mielenkiintoses tilantees ja vaarallises tilantees.” (H4)*

Erään vastaajan esimerkki koski hinnoittelua. Hän totesi, että data-analytiikan avulla osa hinnoittelusta voisi osua oikeaan, osassa taas vaarana olisi, että väärän hinnoittelun takia tuotteet jäisivät myymättä tai niistä saatu kate olisi alhaisempi. Tämän jälkeen vastaaja kuvaili, kuinka tärkeä osuus ihmisellä oli tehdä lopullinen päätös hinnoittelun suhteen, koska ympäristö ja vallitseva tilanne, jota data-analytiikassa oli hankala huomioida, saattoi vaikuttaa merkittävästi lopulliseen hinnanmuodostukseen. Toisen vastaajan tilanne oli samantyyppinen. Hän kuvaili, kuinka esillä pidettiin tuotteita, joita ei oikeastaan myyty juuri ollenkaan. Pelkän data-analytiikan perusteella tällaiset tuotteet karsiintuisivat myynnistä pois. Tosiseikka kuitenkin oli, että nämä ostamatta jäävät tuotteet toimivat eräänlaisena vertailukohteena tuotteilla, joilla taas tehtiin isoin myynti.

Yhdessä organisaatiossa haasteita oli koettu tilanteessa, joka liittyi myös henkilöstöön. Tässä esimerkissä vastaaja kertoi, että aikaisemmin heillä itsenäinen it-osasto vei asioita eri suuntaan, kuin mihin liiketoimintapuolella niitä haluttiin vietävän. Tämä seikka oli kuitenkin kohdeorganisaatiossa ratkaistu organisaatiomuutoksilla, jonka myötä it-osaston toimintaa siirryttiin ohjaamaan päätöksentekijän toimesta.

Puolet vastaajista nimesivät data-analytiikan haasteita ilmenevän tiedon yhdistämisessä. Yksi vastaajista kertoi tiedon yhdistämisen olevan hankalaa, koska tieto oli hajallaan ja eri muodoissa eri järjestelmissä. Kyseisen vastaajan organisaatiossa ongelma oli kuitenkin hyvin tiedostettu ja suunnitteilla oli järjestelmäkehitystä. Tämän vastaajan organisaatiossa ei ollut käytössä BI-ohjelmistoa. Myös toinen vastaaja, jonka käytössä oli vanha BI-ohjelmisto, oli tyytymätön yhdisteltävyyteen ja näki päätöksenteon olevan hankalaa hajallaan olevaan tietoon perustuen. Kolmannella vastaajalla oli BI-ohjelmisto ollut käytössä jo pitkään, mutta tiedon yhdisteltävyydessä oli silti monenlaisia ongelmia. Haastateltava mainitsi ongelmia olevan muun muassa tietojen ajallisessa kohdistamisessa ja luokittelussa. Vastaaja kertoi, että he olivat BI-ohjelmistoa käyttöönottaessa olettaneet, että yhdistely olisi toiminut suoraviivaisemmin. Kokonaiskuvassa tiedon yhdistäminen

vaikutti toimivan, mutta pienissä osissa ja päätöksenteossa virheellisyys saattoi johtaa vääriin tulkintoihin datasta.

*”Kyllä ne teknisesti yhdistyy, mutta ne ei yhdisty loogisesti. Ne ei yhdisty ajallisesti, se logiikka antaa väärän kuvan.” (H4)*

Yksi vastaaja kertoi heillä data-analytiikan haasteita ilmenevän asioiden tulkinnassa eri tahojen välillä. Erityisesti it-osaston ja päätöksentekijän välillä saattoi olla haasteita sen suhteen mitä it pystyi tuottamaan ja mitä päätöksentekijä tarvitsi.

Yhdellä vastaajalla data-analytiikan haasteita ilmeni siinä, miten käyttöoikeudet oli rajattu. Tämän takia vastaaja koki saavansa käyttöönsä usein liian suppeata tietoa. Tämä tulos ei välttämättä ole luokiteltavaksi data-analytiikan haasteeksi, koska ongelma vaikuttaisi muodostuvan enemmänkin siitä, että organisaatiossa ei ole haluttu jakaa tietoa kaikille tahoille. Kuitenkin tämän havainnon perusteella voidaan todeta, että tiedon jakaminen kyseisessä organisaatiossa ei ole ollut avointa ja sen on aiheuttanut haasteita päätöksentekijän päässä. Haastateltava kertoi, että tieto oli ollut enemmän tiettyjen osastojen käytössä, kuin että se olisi ollut koko yrityksen tasolla vapaasti käytettävissä. Tämä toimintatapa olikin organisaatiossa huomattu epäedulliseksi ja tiedon avoimuutta oltiin lähdössä laajentamaan.

*”Nää on vähän kaikki osastot tietyllä tapaa ajautunut vuosien mittaan siihen, et työskennellään vähän siellä omassa siilossa, mut nyt mitä ollaan täs tekemäs justiin ja kun organisoidutaan pikkusen uudestaan, niin ollaan lähetty näitä tämmösiä omia siiloja purkamaan pois.” (H6)*

Yksi vastaajista kuvaili, kuinka heillä oli läpikäyty pitkä prosessi, jonka myötä sisäisen datan laatua oli parannettu. Heillä oli prosessissa keskitytty Master dataan ja sen hallintaan. Hän painotti datan hallinnoinnilla, tietovirtausmallien hahmottamisella, selkeällä rakenteella ja standardeilla olevan suuri merkitys siinä, kuinka organisaation data-analytiikan

raaka-aineesta, datasta, oli saatu yhtenäistä ja vertailukelpoista ja se toimi laadukkaampana tukena päätöksenteolle.

Myös toinen vastaaja kertoi, että datan hyvä laatu ei ollut itsestäänselvyys ja siihen tuli kiinnittää huomiota, jotta organisaatiolle ei aiheutuisi uusia haasteita huonolaatuisen datan hyödyntämisestä päätöksenteossa.

Haastattelujen tekohetkellä COVID-19-pandemia oli vaikuttanut suomessa jo noin 1,5 vuoden ajan. Vallitseva poikkeustilanne nousi haastatteluissa esiin yllättävän vähän. Tähän todennäköisesti vaikutti se, että vallitseva tilanne oli jo monessa organisaatiossa osa arkipäivää. Yhdessä haastattelussa aihe nousi esiin lähinnä positiivisessa hengessä, kun haastateltava kuvaili, kuinka dataa oli voitu hyväksikäyttää tukena päätöksissä, jotka poikkeustilanteen vuoksi olivat olleet hieman normaalista poikkeavia. Toisaalta vastaaja kertoi myös, että poikkeustilanne ei ollut muuttanut heidän data-analytiikan hyödyntämistään vaan sitä oli käytetty samoilla menetelmillä jo aiemminkin.

#### **5.4 Data-analytiikan riittävyys ja tulevaisuuden näkymät**

Kolmannen tutkimuskysymyksen selvittämiseksi haastateltavilta kysyttiin, kaipasivatko he edustamaansa organisaatioon laajempaa data-analytiikkaa ja kuinka he näkivät tulevaisuuden data-analytiikan hyödyntämisen suhteen. Ensimmäiseksi tuloksista selvisi se, koettiin kohdeyrityksissä data-analytiikan tämänhetkinen taso riittävänä. Vastaukset tähän kysymykseen, olivat hyvin samansuuntaisia, kuin aiempaan kysymykseen datan ja datan analysoinnin riittävydestä. Nykytilanne koettiin ihan hyväksi, mutta vastaajat tiedostivat, että kehitystä oli tehtävä. Tässä kysymyksessä korostui aiempaan verrattuna se, että vastaajat vaikuttivat olevan hyvin valveutuneita tilanteestaan ja laajemman data-analytiikan hyödyntämisestä.

*”No kyllä meillä on tulossa ja on käynnissäkin olevia tiettyjä osa-alueita mitä me tiedetään, että mistä meidän tarvii saada enempiä tietoa ja mikä data meiltä*

*puuttuu ja näissä on lähetty nimenomaan sillä kantilla, että meillä on tarve. Me tiedetään, että joo tuo tieto me tarvittais, mutta sitä meillä ei tällä hetkellä oo. Sitä pitää kehittää, jotta me saadaan prosessit kehitettyä, prosesseja vietyä eteenpäin. Suhteellisen paljon meillä on tietoa. Kyllä meiltä vielä puuttuu sitä ja me tarvitaan sitä kyllä lisää, mutta sanotaanko, et järkevällä pohjalla siis sillai, et koko ajan mä sanoin, niin mennään tarkempaan ja tarkempaan, askel kerrallaan, jotta se pystytään myös hyödyntämään se.” (H1)*

Eräs vastaaja kertoi, että heidän organisaatiossaan käytettiin laskennan puolesta hyvin simppeleitä data-analytiikkaa. Tämä saattoikin olla yksi kohde, jonka osalta data-analytiikkaa voitiin laajentaa. Toinen vastaaja kertoi kilpailijoiden tarkemman analyysin olevan heille tärkeä kehityskohde jatkossa. Kolmas vastaaja uskoi ulkoisten datalähteiden käytön lisäämisen olevan askel, jolla tiedonsaanti lisääntyisi. Haastateltava mainitsi hinnoittelun ja investointien olevan omalle liiketoiminnalleen kiinnostavia ulkoisista lähteistä saatavia tietoja.

*”No siis, mun mielestä, laajentamisella mä tarkoitan sitä, että sisäisten tietolähteiden jalostuneella käytöllä, niin täydennetään sit ulkoisten tietolähteiden tiedolla, et pitäis käyttää niitä ulkosiakin tietolähteitä enemmän.” (H5)*

Yksi yhtenevä, lähes kaikissa haastatteluissa pinnalle noussut käsite oli ennustaminen data-analytiikan avulla. Tämä oli mielenkiintoinen havainto, koska huolimatta siitä, että kohdeorganisaatiot toimivat hyvin eri aloilla ja olivat data-analytiikan hyödyntämisen osalta eri vaiheissa, ennustava data-analytiikka nähtiin silti omaan kohdeyritykseen sopivana hyödyntämiskeinona.

*”Meidän raportoinnissa on vielä hyvin pitkälti sitä, että A) ensinnäkin se on hyvin peruutuspeiliä, et katotaan taaksepäin tai sitä, et mikä tilanne on just nyt ja kyl se meidän tavote on, että me päästään siitä eteenpäin semmoseen, et me pystytään entistä enemmän ennustamaan eteenpäin.” (H2)*

Yksi vastaajista kertoi ennustettavuuden olevan tulossa omalle alalleen pikkuhiljaa. Vastaajalla oli kokemusta myös toisenlaisesta alasta ja hän kertoi nähneensä siellä ennustettavuutta käytettävän enemmän. Haastateltava mainitsi omalle alalleen olevan hyödynnettävissä ennustettavuutta useisiin käyttökohteisiin: hinnoitteluun, tilaamiseen, varastohallintaan ja asiakkaan palvelemiseen reaaliaikaisemmin. Hyödyntäminen monissa kohteissa vaati myös toimintojen automatisointia.

*”Tonkin tietää jo entisestä elämästä kaupan alalta, niin että se (ennustettavuus) on ollut olemassa siellä jo pidempään ja sitä ei täällä vielä hyödynnetä.” (H3)*

Toinen vastaaja kommentoi, että ennustamista hyödynnettiin kyllä jossain määrin heidän organisaatiossaan, mutta käytön laajentaminen ja edistyneempien työvälineiden löytäminen oli tarpeen. Nämä tulokset osoittavat, että kohdeorganisaatioissa on tunnistettu ennustavan data-analytiikan tarve ja sen vaikutukset kannattavaan liiketoimintaan. Kuten aikaisemmat tutkimukset osoittavat, ennustavassa data-analytiikassa piilee organisaatioille suuri potentiaali parantaa kilpailukykyä (Gandomi & Haider, 2015; Sircar, 2009, s. 293).

Vain yksi vastaajista ei tuonut ennustavaa data-analytiikkaa itse esiin, joten asiaa tiedusteltiin häneltä erikseen. Vastaaja ei nähnyt sitä niin tärkeäksi omaan työkuvaansa peilaten. Hän kuitenkin mainitsi, että heidän organisaatiossaan esimerkiksi myyntiosastolla ennustavan data-analytiikan käyttö nähtiin paljon tarpeellisempänä.

*”Jos ajatellaan sitä verkkokaupan puolella tai yleensä myynnin puolella, tota ennustettavuutta, niin kyl mä sanoisin, et se tulee siinä isompaan rooliin sitten ja tarpeellisiksi.” (H6)*

Data-analytiikan laajentamisen osalta esille tuli sen tuoma hyöty suhteessa hintaan. Yksi vastaajista kertoi muutamia esimerkkejä, joihin heidän organisaatiossaan data-

analytiikkaa voitaisiin hyödyntää. Joihinkin kohteisiin, jotka oli mahdollista toteuttaa pienillä kustannuksilla, oli tehty pilotteja. Esimerkkinä esiin voidaan nostaa profiloiva asiakasdata. Vastaaja totesi, että heillä data-analytiikan hyödyntämiskohteita oli paljon ja resurssien kohdentamista kuhunkin projektiin oli harkittava tarkoin. Toisen vastaajan kommentissa kiteytyi se, että tuloksia oli vaikea mitata ja sen takia oli hankalaa tehdä päätös, mihin resursseja tulisi suunnata.

*”Mutta noita asioita, kun niitä voidaan niin monesta kantista katkoa, että onko ne tulokset, tuleeko ne heti, tuleeko ne vuosi myöhemmin, markkinointia nyt yleensäkin voidaan, aina voi kyseenalaistaa sillä, että no, ei tämä heti näy kassassa, vaan tämä näkyy pitkällä aikavälillä.” (H3)*

Seuraavaksi haastateltavilta tiedusteltiin, kuinka he näkivät tulevaisuuden data-analytiikan suhteen. Tämän osalta vastaukset olivat vaihtelevia riippuen, katsoiko vastaaja kauas- vai lähitulevaisuuteen. Lähitulevaisuudessa vastaajat kertoivat, että luvassa oli ennustavan data-analytiikan kehitystä, automatiikkaa, edistyneemmän data-analytiikan, koneoppimisen, sensoridatan ja algoritmien hyödyntämistä sekä järjestelmämuutoksia. Ennustavan data-analytiikan osalta yhdelle vastaajalle huolia tuotti se, että sopivaa järjestelmää, joka vastaisi yrityksen tarpeisiin ja soveltuisi laajaan käyttöön, oli vaikea löytää.

*”Mä oon nyt varmaan kattonut 50 eri esitystä asiasta ja en oikeen sellasta, ei sellasta kolahdusta ei oo vielä kertaakaan tullut, että mikä olis se juttu, et okei, tolla mennään.” (H4)*

Muutamit kauemmas tulevaisuuteen katsovat vastaajat mainitsivat tekoälyn, robottiautomaatiikan sekä koneoppimisen olevan osa alan tulevaisuutta. Kaikki vastaajista ei kuitenkaan nähnyt näiden tulevaisuuden trendien vielä hetkeen olevan pinnalla omassa organisaatiossansa.

*”No en oikeestaan täs vaiheessa. Kyllä näihinkin on törmätty ja tarjoajia riittää hyvinkin paljon näihin eri toimintoihin ja robotteihin, mut ei, me ei olla täs vaihees lähdetty ollenkaan oikeestaan millään tavoin vielä kartottamaan näitä, et mitä mahdollisia juttuja olisikaan. Ei nyt aivan lähitulevaisuudessa kumminkaan.” (H6)*

Yksi vastaajista koki henkilökohtaisesti, että edellä mainitut trendit olivat läpikäyneet jonkinlaisen inflaation. Hän kritisoi ajatusta, että esimerkiksi tekoälyä olisi tulossa vain sen takia, että yrityksen strategiana olisi lisätä tekoälyä toimintoihinsa. Hän uskoi, että tekoäly tulisi lisääntymään, mutta motiivit sen hyödyntämiselle tulisi lähteä liiketoiminnan tarpeesta.

Kaksi vastaajaa kertoi heidän organisaatioissaan tekoälyä hyödynnettävän jo pieneltä osin. Näille kohdeyrityksille oli yhteistä, että molemmissa käytettiin ennustamisjärjestelmää, jossa hyödynnettiin tekoälyä. Lisäkysymysten myötä selvisi, että järjestelmän lisäksi toisessa organisaatiossa hyödynnettiin koneoppimista ja toisessa testattiin tekonäköä. Tulevaisuudesta todettiin, että tekoälyn määrän uskottiin lisääntyvän erityisesti tuotannon ja markkinoinnin puolella.

*”Siellä on useita eri projekteja, jossa tulevaisuudessa nään, et semmosia, jotka pysyy sitä manuaalista työtä, tai semmosta bulkkityötä, helpottamaan ja tehostamaan.” (H1)*

Davenport (2018, s. 74–76, 79) tuo tutkimuksessaan esiin tulevaisuuden data-analytiikan suuntauksina myös erityisesti tekoälyn, kuin myös kone- ja syväoppimisen, ohjelmistorobotiikan sekä luonnollisen kielen käsittelyn. Hän esittää nämä edistyksellisemmät data-analytiikan menetelmät jatkumona kehitykselle, jonka organisaatiot ovat jo saavuttaneet kokonaisvaltaisessa data-analytiikan käytössä.

Yksi haastateltavista pohti data-analytiikan tulevaisuutta suhteessa toisiin, alueella toimiviin yrityksiin. Hän näki osaamisen saatavuuden olevan yksi suuri kompastuskivi data-

analytiikan ja yleisesti alueen kehityksen suhteen. Keskustelua syntyi erityisesti it-alan ihmisten hankalasta rekrytointilanteesta, kun kilpailu osaavien ja ammattitaitoisten henkilöiden palkkaamisesta on kiristynyt.

*”Mutta, kyl varmaan nyt tästä eteenpäin, niin alkaa tulla enemmän sitä erkaantumista, että osa menee vauhdikkaasti ja osa jää junnaamaan paikoilleen, että, mut tosi paljon kyllä on kiinni siitä, että tavallaan kuinka paljon alueella saadaan sitä osaamista eteenpäin.” (H2)*

## 5.5 Yhteenveto tuloksista

Haastatteluista saatiin paljon informaatiota käytännöistä aiheeseen liittyen, ja ne syvensivät erityisesti tutkielmantekijän näkemystä tieteellisen tutkimuksen lisäksi reaali-tilanteesta. Tämän tutkielman kannalta, tuloksista relevantteja ovat ne, joiden myötä voidaan vastata tutkielman alussa esitettyihin tutkimuskysymyksiin. Yhteenvedot keskeisistä tuloksista on esitetty seuraavaksi taulukoissa 1–3. Yhteenvedojen tarkoituksena on helpottaa tulosten vertailua sekä osoittaa niissä saavutetut keskeiset erot ja samankaltaisuudet.

Ensimmäisessä yhteenvedossa, taulukko 1:ssä, on esitetty haastateltavien käyttämät datalähteet. Yhteenveto sisältää haastateltavan arvion heillä käytettävän sisäisen ja ulkoisen datan suhteesta, molempien yleisimmät datalähteet sekä esimerkit, minkä tyyppistä dataa haastateltavat mainitsivat hakevansa datalähteistä.

**Taulukko 1.** Datalähteet ja analysointityökalut.

Haastateltava	Datalähteet	Analysointityökalut
---------------	-------------	---------------------

H1	80 % sisäistä dataa: ERP, myynninohjausjärjestelmä (myyntidata). 20 % ulkoista dataa: tavarantoimittajat, tukkuliike	BI-ohjelmisto, ennustamisjärjestelmä, automaattitilausjärjestelmä, Excel, ERP, muut tukijärjestelmät
H2	70 % sisäistä dataa: (asiakas-, transaktio- ja sopimusdata). 30 % ulkoista dataa: viranomaislähteet, rekisterit (asiakas- ja tuotedata)	BI-ohjelmisto, oma räätälöity työkalu, Excel, oma tietokanta, muita visualisointiohjelmia
H3	50 % sisäistä dataa: omat järjestelmät. 50 % ulkoista dataa: internetsivut (tuotedata)	Visualisointiohjelma, myynninohjausjärjestelmä, verkkosivujen kävijäseuranta, oma ja maksetut järjestelmät, Excel
H4	70 % sisäistä dataa: ERP (tuotanto-, talous-, myynti- ja ostodata). 30 % ulkoista dataa: julkiset ja ostetut lähteet, koneet ja laitteet (myynti- ja sensoridata)	BI-ohjelmisto, ennustamisjärjestelmä, oma tietokanta, Excel, muita visualisointiohjelmia
H5	100 % sisäistä dataa: omat järjestelmät (transaktiodata)	Excel, ERP, oma tietokanta, raportointiympäristö
H6	85–90 % sisäistä dataa: ERP (myynti- ja ostodata) 10–15 % ulkoista dataa: viranomaislähteet, rekisterit (asiakasdata)	BI-ohjelmisto, verkkosivujen kävijäseuranta

Kuten yhteenvedosta huomataan, sisäistä dataa käytetään kohdeorganisaatioissa selvästi enemmän ja sitä haetaan ERP:stä tai muista omista järjestelmistä. Eniten haetaan myyntidataa, mutta muutakin käytetään. Ulkoisen datan osalta käyttö on vielä vähäisempää ja datalähteet ovat hajanaisempia eri organisaatioissa. Ulkoisesta datasta tuote- ja asiakasdata ovat kuitenkin haetuimpia. Analysointityökalujen osalta, kaikkien vastaajien osalta tuloksissa toistui se, että analysointityökaluja oli käytössä useita. BI-ohjelmistot sekä Excel oli eniten mainituimpien listalla. Toisaalta muidenkin työkalujen osalta useampi vastaaja nimesi samoja ohjelmia tai työkaluja, kuten ennustamisjärjestelmät, visualisointiohjelmat ja verkkosivujen kävijäseurannat.

Yksi tutkielmassa selvitettävistä seikoista oli, näkivätkö kohdeorganisaatioiden edustajat käyttämänsä datan ja siitä muodostettavan raportoinnin heille olennaiseksi ja riittäväksi, kun asiaa tarkasteltiin liiketoiminnan päätöksenteon kannalta. Tämän osalta havaitut

tulokset on esitetty alla taulukko 2:ssa. Kun kyse oli datan ja sen analysoinnin olennaisuudesta, vastaajat olivat hyvin yksimielisiä siitä, että molemmat olivat heidän päätöksenteossään olennaisia. Datan ja analysoinnin riittävyyden osalta vastaukset olivat edelleen hyvin yhteneväisiä, mutta niistä oli kuitenkin havaittavissa pieniä eroavaisuuksia. Sekä datan, että analysoinnin riittävyys luokiteltiin kohtuullisen riittäväksi, riittäväksi tai melko riittäväksi. Vastauksissa toistui myös pohdinta, että dataa, kuin myös raportointia, oli tarve jatkuvasti kehittää.

**Taulukko 2.** Datan ja analysoinnin olennaisuus ja riittävyys.

Haastateltava	Datan ja analysoinnin olennaisuus	Datan ja analysoinnin riittävyys
H1	On olennaista	On riittävää, jatkuvasti kehitetään
H2	On olennaista	Riittävyys kohtuullisella tasolla, jatkuvasti kehitetään
H3	On olennaista	Riittävyys kohtuullisella tasolla, jatkuvasti kehitetään
H4	On olennaista	Melko riittävää, jatkuvasti kehitetään, koetaan kuitenkin tietynlaisen datan puutetta
H5	On olennaista	On riittävää, kehittämiskohteitakin löytyisi, mutta niihin panostamisen hyödyt hankala osoittaa
H6	On olennaista	Data on riittävää, analysoinnin riittävyys heikompaa, koska data ei yhdisteltävissä

Kolmas haastatteluissa esiin nostettu aihe oli, kaipasivatko yritykset laajempaa data-analytiikkaa ja miten he näkivät tulevaisuuden data-analytiikan saralla. Näihin aiheisiin saatuja tuloksia esitetään taulukko 3:n yhteenvedossa. Ensinnä käsiteltiin data-analytiikan

laajentamista ja tiedusteltiin, mitä kohteita vastaajat havaitsivat organisaatiossaan, joissa data-analytiikasta voitaisiin hyötyä. Eniten esiin noussut vastaus oli, että haastateltavat olivat kiinnostuneet ennustavasta data-analytiikasta ja sen kehittamisestä organisaatiossaan. Puolet vastaajista kertoivat, että heillä oli tulossa järjestelmämuutoksi tai -kehitystä, jonka he uskoivat tuovan muutoksia data-analytiikan hyödyntämiselle. Tulevaisuuden näkyminä vastaajat uskoivat data-analytiikkaa kehitettävän entisestään. Niillä vastaajilla, joilla ei ollut vielä konkreettista strategiaa, kuinka data-analytiikkaa tulotaisiin hyödyntämään, oli kuitenkin näkemyksiä siitä, mihin data-analytiikkaa tulisi heidän organisaatiossaan kohdentaa. Vastauksissa yhtenäistä oli, että kiinnostusta edistyneemmälle data-analytiikalle oli ja muutamassa vastauksessa hyödyntämiskohteiksi mainittiin tekoäly ja koneoppiminen.

**Taulukko 3.** Data-analytiikan laajentaminen ja tulevaisuuden näkymät.

Haastateltava	Data-analytiikan laajentaminen	Tulevaisuuden näkymät
H1	Tarve datalle, data-analytiikkaa pyritään laajentamaan	Tulossa automatisointia, mahdollisesti tekoälyä ja koneoppimista
H2	Tarve datalle ja paremmalle ennustamiselle, järjestelmämuutoksia tulossa	On kiinnostusta edistyneemmälle data-analytiikalle, mutta riippuvaista osaamisesta
H3	Kiinnostusta ennustamiselle ja hinnoittelun automatisoinnille, järjestelmäkehitystä tiedon yhdistämisen helpottamiseksi	On kiinnostusta edistyneemmälle data-analytiikalle, mutta ei vielä konkreettisia toimia
H4	Tarve ennustamiselle, sensoridatan hyödyntäminen	Mahdollisesti tekoälyä, laajamittaisempaa ennustamista
H5	BI-ohjelmisto tulossa, ulkoista dataa käyttöön ja sen avulla ennustamista, koneoppimisalgoritmin testausta	Tarvetta koneoppimiselle ja ohjelmistorobotiikalle

H6	Kiinnostusta tarkemmalle kilpailija-seurannalle, järjestelmämuutoksia tulossa	Tarkempaa seuranta ja suunnittelua
----	---	------------------------------------

## 6 Johtopäätökset

Tässä viimeisessä luvussa esitetään empiirisen tutkielman tuloksista tehtävät johtopäätökset ja verrataan niitä kootusti aikaisempiin tutkimuksiin. Johtopäätösten myötä vastataan tutkielman tavoitteisiin ja alussa asetettuihin tutkimuskysymyksiin. Toisessa alaluvussa otetaan kantaa tutkielman luotettavuuteen ja viimeisessä alaluvussa esitetään ehdotuksia muille mahdollisille jatkotutkimuksille.

### 6.1 Yhteenveto tutkielmasta

Alkuun on vielä hyvä selventää yhteenvetona haastateltavien käyttämä termistö ja sen eroavaisuudet suhteessa tämän tutkielman käsitteisiin. Kuten todettu, tässä tutkielmassa BI esitetään analytiikan varhaisena muotona ja data-analytiikka siitä edistyneempänä menetelmänä. Tätä näkökulmaa tukee esimerkiksi Shardan ja muiden (2018, s. 13) esittämä data-analytiikan elämänkaari. Kaikkien haastateltavien näkemykset ja tulkinnat termeistä eivät olleet yhtenevät tieteellisen tutkimusnäkökulman kanssa. Huolimatta siitä, käyttivätkö vastaajat työorganisaatioissaan jompaakumpaa näistä termeistä tai jotain ihan muuta termistöä, tuloksien esittelyn ja johtopäätösten osalta haastateltavien käyttämät käsitteet on mukautettu tähän tutkimusnäkökulmaan. Tämän perusteella voidaan myös todeta, että tieteellisessä tutkimuksessa sekä käytännön kentällä käytettävä termistö ei ole yhdenmukaista.

Aluksi tutkielmalle asetettiin tavoitteet, joihin vastaamiseksi muodostettiin kolme tutkimuskysymystä. Kun kyse on laadullisesta tutkimuksesta, kysymyksiin ei ilmennyt yhtä vastausta vaan tulkintoja ja kokemuksia ilmeni yhtä monta, kuin haastattelujakin tehtiin. Vastauksista tehdyt havainnot on esitetty tutkielman tulokset -luvussa, ja tässä luvussa tulkitaan kokonaisuutena, mitä enemmistö tutkimuskysymyksiin vastasi. Ensimmäisen tutkimuskysymyksen motiivina oli kartoittaa kohde yritysten käytössä olevaa dataa sekä sen analysointimenetelmiä.

- *Mistä tiedonlähteistä päätöksenteon tueksi kerättävä data koostuu ja minkä avulla sitä analysoidaan?*

Tutkielmassa selvisi mistä tiedonlähteistä päätöksenteon tueksi kerättävä data koostuu. Se koostui enimmäkseen osin haastateltavien organisaatioiden omista järjestelmistä kerätystä strukturoidusta ja sisäisestä datasta. Toisaalta, vaikka ulkoista dataa kerättiin myös yrityksissä vaihtelevasti, sen merkityksen päätöksenteossa nähtiin kasvavan ja voidaan todeta, että sen hyödyntäminen kohdeyrityksissä on kasvusuunnassa. Näin tulosta tukee osittain Kwonin ja muiden (2014, s. 392) tutkimus, jonka mukaan juuri ulkoisen datan käyttö voi tuoda suurempaa hyötyä päätöksenteon tueksi. Kohdeyrityksissä ei ehkä ollut vielä saavutettu ulkoisella datalla tarpeeksi liiketoiminnallista arvoa tai sillä ei vain nähty olevan niin suurta vaikutusta juuri kyseiseen liiketoimintaan.

Tutkielmassa selvisi myös, minkä avulla organisaatioissa analysoidaan kerättyä dataa. Tässä tuloksessa yllättävää oli, kuinka monipuolisesti haastateltavat kertoivat käyttävänsä erilaisia analysointiohjelmiä. Vaikka haastatteluissa nousi esiin samoja ja samankaltaisia työkaluja, siitä huolimatta vastauksissa oli havaittavissa tasoeroja eri yritysten kesken. Siinä missä toinen haastateltava kertoi heidän tilanteensa olevan ihan hyvä data-analytiikan suhteen, toinen pohti, ettei sitä hyödynnetty juuri lainkaan. Edistyneempiä data-analytiikan työkaluja ei vastaajien yrityksissä hyödynnetty yleisesti. Tämä tulos on yhtenevä Nykäsen ja muiden (2016, s. 40) tutkimuksen kanssa, jossa selvisi, että taulukkopohjaiset työkalut ja tiedon visualisointijärjestelmät ovat yrityksissä eniten käytetyimmät työkalut.

Toinen tutkimuskysymys toimi jatkumona ensimmäiselle kysymykselle. Kun ensimmäisessä kysymyksessä selvitettiin, mistä dataa haettiin ja millä sitä analysointiin, toisessa kysymyksessä jatkettiin siitä, kuinka olennaiseksi ja riittäväksi haettu data ja sen analysointi koettiin.

- *Onko tällä hetkellä saatavissa oleva data ja sen analysointi olennaista ja riittävää liiketoiminnan päätöksenteon tueksi?*

Kun kyse oli datan ja sen analysoinnin olennaisuudesta, saatu tulos oli hyvin yksimielinen: vastaajat kokivat, että heidän hyödyntämänsä data ja sen analysointimenetelmät olivat olennaisia heidän työssään. Digitalisoituminen ja datasuuntautunut päätöksenteko ovat varmasti tähän suurimpia vaikuttimia, mutta esimerkiksi Davenportin ja muiden (2001) tutkimus siitä, että datasuuntautunut päätöksenteko on alettu nostamaan mukaan yritysten strategiaan voi myös olla liiketoiminnan puolella yksi selittävä tekijä.

Datan ja sen analysoinnin riittävyyden osalta tulokset olivat hieman hajanaisempia, mutta yhdensuuntainen ajattelumalli oli silti havaittavissa. Pääosin riittävyyden koettiin nykyhetkessä olevan kohtuullisella tasolla. Lönnqvist ja Pirttimäki (2006, s. 38) esittivät tutkimuksessaan liiketoiminnassa haasteita nousevan esiin siinä, mitä yrityksen on tehosta ja hyötypanossuhteeltaan järkevää mitata. Heidän tutkimukseensa verrattuna tutkielman tulos on selitettävissä. Davenport (2013, s. 122) jakoi tutkimuksessaan data-analytiikkapohjaisen päätöksenteon vaiheisiin, joka lähti liikkeelle ongelman tunnistamisesta. Tutkielman tulokset eivät tukeneet sitä, että kohdeyrityksissä olisi päätöksenteko jaettu selkeisiin vaiheisiin, joten tämä myös osaltaan saattoi selittää tuloksien hajanaisuutta. Vastaajat kuitenkin tiedostivat, että tilanne oli alati muuttuva ja jatkuvaa kehitystä sekä datan saatavuuden, että analysointimenetelmien tehokkuuden osalta, tuli suorittaa. Tämä taas osoittaa, että kaikki pyrkivät toiminnassaan eteenpäin, eivätkä halua jäädä kehityksen tai kilpailijoiden jalkoihin. Erot vastauksissa riittävyyden osalta selittyvät myös organisaatioiden erilaisella tilanteella vertailtaessa muihin kohdeyrityksiin. Esimerkiksi yrityksessä, jossa analysointijärjestelmiin oli tulossa muutoksia, ei koettu datan analysointia nykytilanteessa riittävänä. Tätä tulosta tukee myös Davenportin (2006) tutkimus, jonka mukaan yritysten kilpailukyky on kiristynyt ja keinot erottautua ovat vähissä.

Kolmas tutkimuskysymys oli kahteen edelliseen kysymykseen verrattuna itsenäinen. Sen avulla haluttiin selvittää data-analytiikan nykytilannetta yrityksissä, tarvetta edistyneemmälle data-analytiikalle, sekä tulevaisuuden näkymiä sen suhteen.

- *Kaivataanko yrityksissä laajempaa data-analytiikkaa, ja mitkä ovat yritysten tulevaisuuden näkymät data-analytiikan suhteen?*

Tutkielman näkökulma oli asettaa käsitteet BI ja data-analytiikka ikään kuin linjaan, jossa BI on data-analytiikan ensimmäinen muoto ja data-analytiikan käsitteellä kuvaillaan edistyneempää analytiikkaa. Kohdeorganisaatioita tarkasteltiin tästä tutkimusnäkökulmasta. Tarkastelun tuloksena voidaan todeta, että suurin osa yrityksistä on analytiikan suhteen vielä BI-vaiheessa. Toisaalta positiivista on, että data-analytiikkaa hyödynnetään jo kuitenkin monessakin organisaatiossa ja sen käyttö on kasvussa, mutta se ei ole vielä pääsääntöinen hyödyntämismenetelmä.

Vastaus siihen, kaivataanko yrityksissä laajempaa data-analytiikkaa, on monitulkintainen. Vastaajat toivat esiin liiketoimintansa haasteita, jonka kautta tuli ilmi konkreettisia tarpeita, joita yrityksillä oli. Tarpeisiin lukeutui tietynlaisen datan saatavuus, sensoridatan hyödyntäminen, analysointimenetelmien kehittäminen järjestelmämuutosten kautta ja ennustavan data-analytiikan hyödyntäminen. Kaikki edellä mainitut voidaan lukea laajemman data-analytiikan avulla saavutettaviksi toimenpiteiksi. Erityisesti ennustavan data-analytiikan hyödyntäminen oli seikka, jonka useat haastateltavat toivat ilmi lähitulevaisuuden toimenpiteenä. Aikaisemmissa tutkimuksissa Gandomi ja Haider (2015) sekä Sircar (2009, s. 293) mainitsevat juuri ennustavassa data-analytiikassa piilevän suurta potentiaalia kilpailukyvyyn parantamiselle.

Toisaalta vastauksissa toistui myös epävarmuus siitä, minkä toimenpiteiden toteuttaminen olisivat liiketoiminnalle arvokkaita. On huomattavaa, että kaikki haastatellut yritykset ovat hyvää tulosta tekeviä suuryrityksiä. Niinpä esimerkiksi niissä organisaatioissa, joissa data-analytiikkaa hyödynnetään vähiten, sen käytöllä ei ole vielä nähty niin suurta

merkitystä juuri kyseiselle liiketoiminnalle. Vaikka laajemman data-analytiikan mahdollisuuksia omassa organisaatiossa tunnistettiin, niiden vaikutukset toiminnan tehostamiseen ja sitä kautta positiiviseen vaikutukseen toiminnan tuloksessa, koettiin vaikeaksi todistaa. Yhteenvetona voidaan esittää, että kohteita laajemmalle data-analytiikalle oli, mutta varsinainen tarve sen hyödyntämiselle ei ollut yhtä vahva.

Tulevaisuuden suhteen vastaajat olivat yhtä mieltä: data-analytiikkaa tulotaisiin varmasti hyödyntämään. Aikajana saattoi joillain olla hyvinkin lähitulevaisuudessa, kun taas toiset kokivat suuremman kehityksen olevan vielä kauempana tulevaisuudessa. Kiinnostusta edistyneemmän data-analytiikan käytölle löytyi ja osa vastaajista mainitsi myös konkreettisia toimia, johon edistyneempää data-analytiikkaa voitaisiin hyödyntää. Tekoäly ja sen esiasteet, kuten koneoppiminen, mainittiin mahdollisina hyödyntämiskohteina. Tulosta tukee Davenportin (2018, s. 74–76, 79) tutkimus, jossa hän nimeää koneoppimisen ja erityisesti tekoälyn mullistavimmaksi teknologiseksi suuntaukseksi.

Tutkielman kohteena olleet yritykset valikoitiin niiden koon ja sijainnin perusteella. Yritykset eivät muutoin olleet vertailtavissa toisiinsa esimerkiksi toimialan tai infrastruktuurinsa perusteella. Yritysten tarpeet olivat erilaisia ja lisäksi ne olivat eri vaiheissa, mitä tuli data-analytiikan hyödyntämiseen. Data-analytiikka on kuitenkin universaali menetelmä ja yritykset toimialasta riippumatta voivat hyödyntää sitä päätöksenteossaan. Niinpä tämän tutkielman yksi kohokohta oli huomata, että yrityksistä löytyi paljon yhteneväisyyttä, mitä tuli BI:n tai data-analytiikan hyödyntämiseen. Yksittäisiä esimerkkejä, erilaisia tilanteita ja haasteita omassa liiketoiminnassa oli paljon, mutta silti havaittavissa oli samankaltaisuutta ja yhtenäistä suuntautumista. Tutkielman kenties suurin oivallus oli tulosten perusteella havainnoida, mikä BI:n ja data-analytiikan tila on eteläpohjalaisissa yrityksissä. Voidaan todeta, että suurin osa on analytiikan käytössä BI-tasolla, mutta joukkoon mahtui myös muutamia, joissa data-analytiikan hyödyntäminen oli jo suhteessa hieman edistyneellisempää.

Lopuksi yhteenvetona voidaan todeta, että yrityksissä koettiin heidän olevan hyvällä tasolla data-analytiikan suhteen. Vaikka kaikki haastateltavat uskoivat, että data-analytiikkaa kehitetään myös tulevaisuudessa, sillä ei silti uskottu olevan niin suurta lisäarvoa liiketoiminnan päätöksenteossa, kuin aikaisempien tutkimusten valossa sillä voitaisiin todeta olevan. Vastauksien perusteella välittyi tieto, että nykyinen liiketoiminta koettiin sen verran vahvaksi ja tehostetuksi, ettei laajemmalla data-analytiikalla uskottu saavutettavan riittävän suuria hyötyjä. Toisaalta juurisyytä näihin ajatuksiin voi olla useita. Yksi mahdollinen selitys on, että data-analytiikan osaaminen ja edistyksellisten menetelmien, kuten laskentatekniikoiden, ymmärrys, ei ole riittävän vahvaa, jotta niiden tuoma lisäarvopotentiaali tunnistettaisiin. Toinen syy voi olla, ettei kaikkia liiketoiminnan piileviä mahdollisuuksia ole vielä löytynyt.

## 6.2 Tutkielman luotettavuus

Tutkielman luotettavuuden arvioinnissa on syytä huomioida muutamia tälle tutkielmalle ominaisia seikkoja. Tutkielma suoritettiin haastattelemalla kuutta liiketoiminnan johtotai päällikkötasossa olevaa henkilöä. Näin pienen otannan perusteella tutkielma ei ole yleistettävä tutkimus Etelä-Pohjanmaan tilanteesta. Voidaan kuitenkin todeta, että tutkielman tavoitteiden saavuttamiseksi haastattelujen määrä on riittävä. Kun myös huomioidaan käytetyt tutkimusmenetelmät, voidaan todeta, että aineiston laatu on pidetty hyvänä, sekä valikoimalla tutkielman tarkoitukseen sopivia haastateltavia, että keräämällä laadukasta aineistoa teemahaastattelujen avulla.

Tutkielman luotettavuutta voidaan arvioida monella eri menetelmällä. Tieteellisessä tutkimuksessa arviointi jaetaan usein kahteen eri menetelmään: tutkimuksen reliaabeliukseen sekä validiukseen. Hirsjärven ja muiden (2007) mukaan reliaabelius kuvaa tulosten toistettavuutta jonkun toisen suorittamana. Validiusta he kuvaavat tutkimusmenetelmän kykyä soveltua mittaamaan tarkoitettua tietoa. Tutkimusvaiheessa huomiota on kiinnitetty erityisesti siihen, että haastateltavan sanoma on saatu esiin sellaisena, kuin hän on sen tarkoittanut. Teemahaastattelun luonne ja sen suomat haastattelukohtaiset

lisä- ja tarkentavat kysymykset ovat olleet hyödyksi luotettavan kokonaiskuvan luomisessa. Tämän perusteella voidaan todeta, että vastaava tutkielma toisen tutkijan toteuttamana antaisi todennäköisesti varsin samanlaisen tuloksen eli reliaaabelius on todettavissa. Toki huomioitavaa on, että vastaukset on kerätty haastateltavien subjektiivisista näkemyksistä ja eri vastaajien tutkiminen saattaisi antaa erilaisen tuloksen. Vastaava tutkielma toistettaessa, olisi luonteeltaan kuitenkin aina uusi tapauskohtainen tutkielma.

Teemakysymykset toimitettiin haastateltaville ennakoon, jolloin heillä oli mahdollisuus valmistautua haastattelutilanteeseen. Tulosten analysointivaiheessa aineistona huomiointiin koko haastatteluaineisto, myös vastaajien tuomat uudet näkemykset ja lisähuomiot aihetta sivuten. Näiden seikkojen ja käytetyn tutkimusmenetelmän luonteen avulla todettiin, että myös tutkielman validius saavutettiin.

Tutkielman tuloksissa ilmeni seikkoja, jotka olivat yhteneviä aikaisempien tutkimusten kanssa. Tämä nostaa tutkielman luotettavuutta. Toisaalta tuloksissa ilmeni myös seikkoja, jotka erosivat teoreettisesta tutkimuksesta. Tämä taas osoittaa, että tulosten tulkinta on ollut tutkijan omista ennakoasetelmista riippumatonta. Myös haastateltavien vertailussa esiin nousi paljon samankaltaisia vastauksia ja näkemyksiä, joten tämäkin seikka tuo tutkielman tavoitteille luotettavuutta. Tutkielmassa saatiin mitattua juuri niitä asioita, joihin tutkimuskysymysten avulla ja tutkielman todellisten tavoitteiden myötä haluttiin saada vastaus.

### **6.3 Jatkotutkimusehdotukset**

Päätöksenteko organisaatioissa pohjautuu yhä enemmän dataan. BI sekä data-analytiikka aiheina ovat ajankohtaisia nyt ja tulevaisuudessa. Tosin niistä käytettävät termit saattavat hyvinkin pian muuttua ja teknologiakehitys sekä tekoäly tuoda tullessaan uusia menetelmiä. Tämän tutkielman teoreettinen viitekehys antaa lukijalle hyvän pohjatiedon käsitelystä aihepiiristä, ja tulosten pohjalta voi herätä uusia tutkimusideoita. Myös tutkijan tietämys aiheesta parani merkittävästi ja antoi eväitä ammatillisen osaamisen

kehittämiseen. Alla esitetään esimerkkejä tutkimuksista, joita tutkielmassa esitettyjen teemojen pohjalta voisi suorittaa.

Tutkielma suoritettiin Etelä-Pohjanmaan alueella, joten vastaavia tutkimuksia voisi tehdä muuallakin Suomessa. Tutkimus voitaisiin suorittaa myös kyselytutkimuksena, jolloin vastauksista voitaisiin johtaa yleistyksiä tutkitun alueen data-analytiikan tilasta. Tutkielma suoritettiin eri ympäristöissä toimiville yrityksille, joten tutkimusta voisi kohdentaa tietyille toimialalle tai yrityksiin, joissa olisi keskenään samansisältöistä rakennetta, esimerkiksi tuotantolaitoksiin. Tutkielma suoritettiin myös haastattelemalla eri liiketoimintojen päällikkö- ja johtotason henkilöstöä. Tämänkin osalta jatkotutkimuksissa voisi keskittyä ennalta määritetyn osaston johtohenkilöstöön. Mielenkiintoista olisi kohdistaa tutkimus esimerkiksi markkinoinnin johdon näkemykseen data-analytiikasta.

Jatkotutkimuksissa voisi perehtyä yrityksiin, joissa käytetään nimenomaan BI:ä, koska tieteellisen tutkimuksen näkemys ja termin käytäntö yritysmaailmassa voivat poiketa toisistaan huomattavasti. Myös data-analytiikkaa hyödyntäviä yrityksiä voisi tutkia edelleen käyttäen kohteena yrityksiä, joissa on käytössä samansuuntaisia menetelmiä, kuten enustavaa data-analytiikkaa. Myös tutkielmassa esitetyt tulevaisuuden suuntaukset data-analytiikan saralla, kuten tekoäly ja sen eri alamuodot, ovat enenevässä määrin käytössä yrityksissä. Näiden osalta olisi mielenkiintoista selvittää tarkemmin mikä käytännön hyödyntämistaso on, esimerkiksi juuri Etelä-Pohjanmaalla.

## Lähteet

- Acito, F. & Khatri, V. (2014). *Business analytics: Why now and what next?* Business Horizons. Vol. 57, 565–570. <https://doi.org/10.1016/j.bushor.2014.06.001>
- Alasuutari, P. (2011). *Laadullinen tutkimus 2.0*. Tampere:Vastapaino.
- Appelbaum, D., Kogan, A., Vasarhelyi, M. & Yan, Z. (2017). *Impact of business analytics and enterprise systems on managerial accounting*. International Journal of Accounting Information Systems. Vol. 25, 29–44. <https://doi.org/10.1016/j.acinf.2017.03.003>
- Bayrak, T. (2015). *A review of business analytics: A business enabler or another passing fad*. Procedia – Social and behavioral sciences, Vol. 195, 230–239. <https://doi.org/10.1016/j.sbspro.2015.06.354>
- Begum, M. S. & George, A. (2017). *A Survey on Data Analytics Framework*. International Journal of Engineering and Technology. Vol. 9, No. 3, 1650–1656. <https://doi.org/10.21817/ijet/2017/v9i3/170903010>
- Bhatt, G. (2001). *Knowledge management in organizations: examining the interaction between technologies, techniques, and people*. Journal of Knowledge Management. Vol. 5, No. 1, 68–75. <https://doi.org.proxy.uwasa.fi/10.1108/13673270110384419>
- Chen, H., Chiang, R. & Storey, V. (2012). *Business intelligence and analytics: From big data to big impact*. MIS Quarterly, Vol. 36, No. 4, 1165–1188. <https://doi.org/10.2307/41703503>
- Davenport, T. & Harris, J. (2007). *Analysoi ja voita: kilpailun uusi tiede*. Talentum Helsinki.
- Davenport, T., Harris, J., De Long, D. & Jacobson, A. (2001). *Data to Knowledge to Results: Building an analytic capability*. California Management Review. Vol. 43, No. 2, 117–138. Noudettu 20.7.2021 osoitteesta: <https://www-proquest-com.proxy.uwasa.fi/docview/216132333/abstract/C7FBD1DC6FC4686PQ/1?accountid=14797>
- Davenport, T. (2006). *Competing on Analytics*. Harvard Business Review. Vol. 84, 98–107. Noudettu 18.7.2021 osoitteesta:

- <http://web.a.ebscohost.com.proxy.uwasa.fi/ehost/pdfviewer/pdfviewer?vid=2&sid=10465fd5-6a94-47d9-b1b2-dd8e8a462fb7%40sessionmgr4006>
- Davenport, T. (2013). *Keep Up with Your Quants*. Harvard Business Review. Vol. 91, No. 7/8, 120–123. Noudettu 20.7.2021 osoitteesta: <https://www-proquest-com.proxy.uwasa.fi/docview/1371845190/?pq-origsite=primo>
- Davenport, T. (2018). *From analytics to artificial intelligence*. Journal of Business Analytics. Vol. 1, No. 2, 73–80. <https://doi.org/10.1080/2573234X.2018.1543535>
- Digia (2021). *Data ja analytiikka*. Noudettu 10.7.2021 osoitteesta: <https://digia.com/palvelumme/data-ja-analytiikka/>
- Gamage, P. (2016). *Big Data: are accounting educators ready?* Accounting and Management Information Systems. Vol. 15, No. 3, 588–604. Noudettu 20.7.2021 osoitteesta: <https://www-proquest-com.proxy.uwasa.fi/docview/1930956693/abstract/B8CA1884F31B4A87PQ/1?accountid=14797>
- Gandomi, A. & Haider, M. (2015). *Beyond the hype: big data concepts, methods, and analytics*. International Journal of Information Management. Vol. 35, No 2, 137–144. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- Gartner (2022). *Information Technology Gartner Glossary*. Noudettu 31.1.2021 osoitteesta: <http://www.gartner.com/it-glossary/big-data/>.
- Griffin, R. (2012). *Using big data enterprise to combat fraud*. Financial Executive. Vol. 28, No. 10, 44–47. Noudettu 20.7.2021 osoitteesta: <https://www-proquest-com.proxy.uwasa.fi/docview/1266382842/abstract/5B57F2B9981242BFPO/1?accountid=14797>
- Hirsjärvi, S. & Hurme, H. (2008). *Tutkimushaastattelu: teemahaastattelun teoria ja käytäntö*. Helsinki: Gaudeamus.
- Hirsjärvi, S., Remes, P. & Sajavaara, P. (2007). *Tutki ja kirjoita*. Otavan Kirjapaino Oy, Keuruu.
- Isik, Ö., Jones, M. & Sidorova, A. (2013). *Business intelligence success: The roles of BI capabilities and decision environments*. Information & Management. Vol. 50, No. 1, 13–23. <https://doi.org/10.1016/j.im.2012.12.001>

- Janssen, M., van der Voort, H. & Wahyudi, A. (2017). *Factors influencing big data decision-making quality*. Journal of Business Research, Vol. 70, 338–345. <http://dx.doi.org/10.1016/j.jbusres.2016.08.007>
- Khade, A. A. (2016). *Performing Customer Behavior Analysis using Big Data Analytics*. Procedia Computer Science Vol. 79, 986–992. <https://doi.org/10.1016/j.procs.2016.03.125>
- Kirjanpitolaki* 30.12.1997/1336. Finlex. Noudettu 28.3.2022 osoitteesta <https://www.finlex.fi/fi/laki/ajantasa/1997/19971336>
- Kiron, D., Kirk Prentice, P. & Boucher Ferguson, R. (2012). *Innovating with analytics*. MIT Sloan Management Review. Vol. 54, No. 1, 47–51. Noudettu 20.7.2021 osoitteesta: <https://www-proquest-com.proxy.uwasa.fi/docview/1115278032/abstract/DC975FF1BCDC48EFPQ/1?accountid=14797>
- Kowalczyk, M. & Buxmann, P. (2015). *An ambidextrous perspective on business intelligence and analytics support in decision processes: insights from a multiple case study*. Decision Support Systems. Vol. 80, 1–13. <https://doi.org/10.1016/j.dss.2015.08.010>
- Kumar, E. B. & Thiagarasu, V. (2019). *Big Data and its Applications: A Review*. International Journal of Intelligent Computing and Technology. Vol. 2, No. 2, 1–10. Noudettu 20.2.2022 osoitteesta: <https://ijict.com/V2I2/V2I2P01.pdf>
- Kwon, O., Lee, N. & Shin, B. (2014). *Data quality management, data usage experience and acquisition intention of big data analytics*. International Journal of Information Management. Vol. 34, No. 3, 387–394. <https://doi.org/10.1016/j.ijinfomgt.2014.02.002>
- Laihonen, H., Hannula, M., Helander, N., Ilvonen, I., Jussila, J., Kukko, M., Kärkkäinen, H., Lönnqvist, A., Myllärniemi, J., Pekkola, S., Virtanen, P., Vuori, V. & Yliniemi, T. (2013). *Tietojohdaminen*. Tampereen teknillinen yliopisto: Tietojohdamisen tutkimuskeskus Novi.
- Lawless, M. (2014). *Predictive Analytics: An Opportunity for Better Demand Planning and Forecasting*. Journal of Business Forecasting. Vol. 33, No. 4, 44–46. Noudettu 7.8.2021 osoitteesta: <https://www-proquest-com.proxy.uwasa.fi/scholarly->

journals/predictive-analytics-opportunity-better-demand/docview/1657419233/se-2?accountid=14797

- Lindholm, M. (2021, 10. helmikuuta). *Dataräjähdyks muutti datan merkityksen yhteiskunnassa – miten siihen pitäisi reagoida?* Noudettu 9.8.2021 osoitteesta: <https://www.stat.fi/tietotrendit/blogit/2021/datarajahdys-muutti-datan-merkityksen-yhteiskunnassa-miten-siihen-pitaisi-reagoida/>
- Lönnqvist, A. & Pirttimäki, V. (2006). *The measurement of Business Intelligence*. Information Systems Management, Vol. 23, No. 1, 32–40. <https://doi.org/10.1201/1078.10580530/45769.23.1.20061201/91770.4>
- Maisel, L. & Cokins G. (2014). *Why analytics will be the next competitive edge*. The Journal of Corporate Accounting & Finance, Vol. 25, No. 5, 63–72. <https://doi.org.proxy.uwasa.fi/10.1002/jcaf.21974>
- Markkula, T. & Syväniemi, A. (2015). *Analytiikkamatka datasta tietoon ja tiedolla johtamiseen*. Saarijärven Offset Oy.
- Mazzuto, G. & Ciarapica, F. E. (2019). *A big data analytics approach to quality, reliability, and risk management*. International Journal of Quality & Reliability Management. Vol. 36, No. 1, 2–6. <https://doi.org/10.1108/IJQRM-01-2019-294>
- McKinsey Global Institute. (2011). *Big data: The next frontier for innovation, competition, and productivity*. Noudettu 11.7.2021 osoitteesta: [https://www.mckinsey.com/~media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI\\_big\\_data\\_exec\\_summary.pdf](https://www.mckinsey.com/~media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_exec_summary.pdf)
- Mirarab, A., Mirtaheri, S.L. & Asghari, S. A. (2019). *Value creation with big data analytics for enterprises: a survey*. Telkomnika. Vol. 17, No. 6, 2790–2802. <https://doi.org/10.12928/TELKOMNIKA.v17i6.11962>
- Nuopponen, A. (2020). Systemaattinen käsiteanalyysi tutkijan työssä. Teoksessa H. Katajamäki (toim.), *Tieteellinen kirjoittaminen tiedeyhteisössä* (s. 94–122). Vakki ry. Noudettu 24.6.2021 osoitteesta: [https://osuva.uwasa.fi/bitstream/handle/10024/11441/Osuva\\_Tieteellinen\\_kirjoittaminen\\_tiedeyhteis%c3%b6ss%c3%a4\\_2020.pdf?sequence=5&isAllowed=y](https://osuva.uwasa.fi/bitstream/handle/10024/11441/Osuva_Tieteellinen_kirjoittaminen_tiedeyhteis%c3%b6ss%c3%a4_2020.pdf?sequence=5&isAllowed=y)

- Nykanen, E., Järvenpää, M. & Teittinen, H. (2016). *Business intelligence in decision making in Finnish enterprises*. Nordic journal of business. Vol. 65, No. 2, 24–44. Noudettu 27.8.2021 osoitteesta: [http://njb.fi/wp-content/uploads/2016/09/Nykanen\\_Jarvenpaa\\_Teittinen.pdf](http://njb.fi/wp-content/uploads/2016/09/Nykanen_Jarvenpaa_Teittinen.pdf)
- Olszak, C. (2016). *Toward better understanding and use of business intelligence in organisations*. Information systems management. Vol. 33, No. 2, 105–123. <https://doi.org/10.1080/10580530.2016.1155946>
- Pirttimäki V. (2007). *Business Intelligence as a Managerial Tool in Large Finnish Companies*. Tampereen teknillinen yliopisto. Julkaisu 646. Tampere. <http://urn.fi/URN:NBN:fi:ttty-200810021126>
- Puusa, A., Juuti, P., & Aaltio, I. (2020). *Laadullisen tutkimuksen näkökulmat ja menetelmät*. Gaudeamus Oy.
- Radaelli, C. (2000). *Policy transfer in the european union: institutional isomorphism as a source of legitimacy*. Governance. Vol. 13, No. 1, 25–43. <https://doi.org/10.1111/0952-1895.00122>
- Rikhardsson, P. & Yigitbasioglu, O. (2018). *Business intelligence & analytics in management accounting research: Status and future focus*. International Journal of Accounting Information Systems. Vol. 29, 37–58. <https://doi.org/10.1016/j.accinf.2018.03.001>
- Salo, I. (2013). *Big Data: Tiedon vallankumous*. Docendo Oy.
- Schmarzo, B. & Partlow, D. (2013). *Big data: Understanding how data powers big business*. John Wiley & Sons, Incorporated.
- Sharda, R., Delen, D. & Turban, E. (2018). *Business intelligence, analytics, and data science: A Managerial Perspective*. Boston: Pearson.
- Sharma, R., Mithas, S. & Kankanhalli, A. (2014). *Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations*. European Journal of Information Systems. Vol. 23, No. 4, 433–441. <https://doi.org/10.1057/ejis.2014.17>

- Shollo, A. & Galliers, R. (2016). *Towards an understanding of the role of business intelligence systems in organisational knowing*. Information Systems Journal. Vol. 26, No. 4, 339–367. <https://doi.org/10.1111/isj.12071>
- Sircar, S. (2009). *Business intelligence in business curriculum*. Communications of the association for information systems. Vol. 24, No. 17, 289–302. <https://doi.org/10.17705/1cais.02417>
- Soraya, S. (2018). *Data Analytics and Big Data*. John Wiley & Sons, Incorporated.
- TietoEvry (2020). *Ota datamaailman termit haltuun datasanaston avulla*. Noudettu 19.5.2021 osoitteesta: <https://www.tietoevry.com/fi/uutishuone/kaikki-uutiset-ja-tiedotteet/blogi/2020/ota-datamaailman-termit-haltuun-datasanaston-avulla/>.
- Torres, R., Sidorova, A. & Jones, M. (2018). *Enabling firm performance through business intelligence and analytics: a dynamic capabilities perspective*. Information & Management. Vol. 55, 822–839. <https://doi.org/10.1016/j.im.2018.03.010>
- Vera-Baquero, A., Colomo Palacios, R., Stantchev, V. & Molloy, O. (2015). *Leveraging big-data for business process analytics*. The Learning Organization. Vol. 22, No. 4, 215–228. <https://doi.org/10.1108/TLO-05-2014-0023>
- Viitala, R. & Jylhä, E. (2019). *Johtaminen: keskeiset käsitteet, teorit ja trendit*. Helsinki: Edita.
- Väre, T. (2019). *Master Data*. Alma Talent Helsinki.
- Watson, H. J. (2015). *Business Analytics Insight: Hype or Here to Stay?* Business Intelligence Journal. Student Edition, 33–37. Noudettu 20.7.2021 osoitteesta: <https://www-proquest-com.proxy.uwasa.fi/docview/1676462842/abstract/8973F414E8394CBCPQ/1?accountid=14797>
- Watson, H. J. & Wixom, B. H. (2007). *The current state of business intelligence*. Computer. Vol. 40, No. 9, 96–99. <https://doi.org/10.1109/MC.2007.331>

## Liitteet

### LIITE 1. Teemahaastatteluiden kysymykset

#### Taustatiedot

- Yrityksen nimi, päätoimiala, liikevaihto, työntekijöiden määrä?
- Haastateltavan nimi, titteli, tehtävä organisaatiossa ja työnkuva?
- Työkokemus alalla ja työkokemus kohdeyhtiössä?
- Mitä ajattelet termeistä BI ja data-analytiikka?
- Onko teillä käytössä jompikumpi termeistä BI tai data-analytiikka?

#### Mistä tiedonlähteistä päätöksenteon tueksi kerättävä data koostuu ja minkä avulla sitä analysoidaan?

- Mistä sisäisistä ja ulkoisista tiedonlähteistä keräätte dataa?
- Prosentuaalisesti paljonko käyttämästänne datasta on sisäistä ja paljonko ulkoista?
- Miten ja millä työkaluilla analysoitte dataa?
- Onko BI-ohjelmisto helpottanut tiedon yhdistelyä?

#### Onko tällä hetkellä saatavissa oleva data ja sen analysointi olennaista ja riittävää liiketoiminnan päätöksenteon tueksi?

- Onko käyttämänne data ja raportointi olennaista?
- Onko kaikki saatavilla oleva data ja raportointi tarpeellista?
- Tulisiko raportointia olla enemmän?
- Onko käyttämänne data ja raportointi riittävää?
- Onko data ja raportointi laadukasta ja luotettavaa?
- Miten päätöksentekoprosessi on kartoitettu?
- Kuinka paljon päätöksenteosta on dataperusteista ja paljonko se perustuu mutuntumaan?
- Mitkä osastot käyttävät data-analytiikkaa?

**Liiketoiminnan ja data-analytiikan haasteet**

- Millaisia haasteita teillä on liiketoiminnassa?
- Onko teillä ollut haasteita data-analytiikan käytön kanssa?
- Oletteko saaneet data-analytiikasta apua haasteiden selättämiseen?

**Kaivataanko yrityksissä laajempaa data-analytiikkaa, ja mitkä ovat yritysten tulevaisuuden näkymät data-analytiikan suhteen?**

- Onko data-analytiikan käytön taso riittävä?
- Koetteko, että yrityksessänne käytetään data-analytiikkaa laajasti?
- Oletteko kartoittanut tarvetta laajentaa data-analytiikan käyttöä?
- Millaiset tulevaisuuden näkymät yrityksellänne on data-analytiikalle?

**Lopuksi**

- Onko jotain muuta aiheeseen liittyvää, mitä haluaisit vielä tuoda esiin?