



**LUT School of Business and Management**

Kauppätieteiden kandidaatintutkielma

Talousjohtaminen

**Datan ja analytiikan hyödyntämisen vaikutus  
työntekijän suorituskykyyn**

The Impact of the Utilization of Data Analytics on Employee-level  
Work Performance

16.1.2019

Maria Häärä

Ohjaaja: Henri Hussinki

## TIIVISTELMÄ

<b>Tekijä:</b>	Maria Häärä
<b>Tutkielman nimi:</b>	Datan ja analytiikan hyödyntämisen vaikutus työntekijän suorituskykyyn
<b>Akateeminen yksikkö:</b>	School of Business and Management
<b>Koulutusohjelma:</b>	Kauppätiede / talousjohtaminen
<b>Ohjaaja:</b>	Henri Hussinki
<b>Hakusanat:</b>	Data, data-analytiikka, työntekijän suorituskyky, regressioanalyysi

Data-analytiikan hyödyntäminen, eli yhtiön keräämän datan jalostaminen päätöksenteon tueksi on noussut organisaatioiden tärkeäksi kilpailuedun lähteeksi. Analytiikan hyödyntämistä onkin siis syytä tutkia sen ajankohtaisuuden ja viimeaikaisen kasvun vuoksi. Tällä tutkielmalla pyritään selvittämään erään energia-alan yhtiön talousyksiköiden datan ja data-analytiikan hyödyntämisen vaikutusta yksittäisen työntekijän suorituskykyyn. Aihetta tutkitaan tutkimusmallin kautta, joka jakaa aiempien tutkimusten pohjalta datan ja analytiikan hyödyntämisen analyttisiin taitoihin, teknologisiin kyvykkyyksiin, datalähtöiseen toimintakulttuuriin sekä organisaatiotasoiseen oppimiseen.

Tutkimusmenetelmänä käytetään lineaarista regressioanalyysia, jonka avulla tarkastellaan, onko tutkimusmallin neljällä osa-alueella ja työntekijän suorituskyvyllä tilastollisesti merkittävää yhteyttä. Tutkimuksessa hyödynnetty aineisto kerättiin kyselylomakkeella, joka lähetettiin tutkittavan yhtiön talousyksiköihin.

Tutkimustulokset osoittavat yhteyden datan ja analytiikan hyödyntämisen sekä työntekijän suorituskyvyn välille. Regressioanalyysin perusteella kuitenkin vain analyttisillä taidoilla on tilastollisesti merkittävä yhteys yksilön suorituskykyyn. Teknologisilla kyvykkyyksillä, datalähtöisellä toimintakulttuurilla ja organisaatiotasoisella oppimisella yhteyttä ei vaikuta olevan. Tutkimustulokset eivät kuitenkaan ole täysin luotettavia eivätkä yksiselitteisiä, koska tutkittava aineisto muodostui erittäin pieneksi ja tutkimus suoritettiin vain yhdelle organisaatiolle.

## ABSTRACT

**Author:** Maria Häärä  
**Title:** The Impact of the Utilization of Data Analytics on Employee-level Work Performance  
**School:** School of Business and Management  
**Degree programme:** Business Administration / Financial Management  
**Supervisor:** Henri Hussinki  
**Keywords:** Data, data analytics, employee-level work performance, regression analysis

Data analytics utilization has become a source of competitive advantage for organizations. Companies use analytics to help decision making by processing data. Because the utilization of data-analytics is topical and it has recently grown, it is important to examine the subject. This thesis pursues to find out the impact of the utilization of data analytics on employee-level work performance in an energy company's Finance Units. The research model was compiled from previous studies and it divides data analytics utilization into four parts: analytical skills, technology capabilities, data-driven culture and organizational learning.

Linear regression analysis was used to examine the relationship between data analytics and employee performance. The data used in the study was collected with questionnaire from the company's Finance Units.

The results confirm the hypothesis on the positive effect of data analytics utilization on employee performance. However, according the regression analysis, statistically significant correlation appears only between analytical skills and employee performance. The study could not confirm direct effect between other variables. The findings are not completely generalizable, as the number of observations was quite small, and the phenomenon was studied in only one organization.

# SISÄLLYSLUETTELO

<b>1 JOHDANTO</b> .....	1
1.1 Tutkimuksen tavoitteet, rajaukset ja tutkimuskysymys .....	2
<b>2 TEOREETTINEN TAUSTA</b> .....	3
2.1 Data ja data-analytiikka .....	3
2.2 Tietoon perustuva päätöksenteko .....	5
2.2.1 Datalähtöinen toimintakulttuuri .....	6
2.3 Suorituskyky .....	7
<b>3 TUTKIMUSMALLI JA HYPOTEESIT</b> .....	9
<b>4 TUTKIMUSMENETELMÄ JA -AINEISTO</b> .....	12
4.1 Analyysimenetelmä .....	12
4.2 Aineiston keräys ja kyselylomakkeen muodostaminen .....	14
<b>5 TUTKIMUSTULOKSET</b> .....	16
5.1 Aineiston kuvailu .....	16
5.2 Faktorianalyysi .....	19
5.2.1 Summamuuttujien muodostaminen faktorianalyysin pohjalta .....	23
5.3 Lineaarinen regressioanalyysi .....	25
5.4 Tutkimustulosten luotettavuus .....	29
<b>6 YHTEENVETO JA JOHTOPÄÄTÖKSET</b> .....	31
<b>LÄHDELUETTELO</b> .....	34

## LIITTEET

Liite 1. Kyselylomake

Liite 2. Summamuuttujien jakaumat

Liite 3. Yksittäisten väittämien perustiedot

Liite 4. Residuaalikuvaajat

## 1 JOHDANTO

Analytiikka on muuttanut perinteistä organisaatioiden toimintaa. Ennen päätökset tehtiin intuition ja tunteen kautta, mutta nyt analytiikan ansiosta päätöksenteossa voidaan tukeutua konkreettiseen tietoon. (Gupta 2016, 1) Yhä useammat organisaatiot ovat tulleet tietoisiksi tiedon analysoinnin tuomista mahdollisuuksista kilpailun saralla, joten yritykset ovat alkaneet hyödyntää edistynyttä analytiikkaa toiminnassaan ja strategiassaan (Davenport & Harris 2007, 227). Analytiikan kasvun räjähdysten ovat mahdollistaneet muun muassa edullinen datan tallennus sekä korkealaatuinen tietojen käsittely (Gupta 2016, 2).

Salo (2014, 6) vertaa datan tuomaa murrosta teolliseen vallankumoukseen, koska datasta on muodostumassa kaikkeen toimintaan tarvittava voimavara ja raaka-aine. Kaikki organisaatiot tarvitsevat dataa ainakin jossain määrin, ja toiset organisaatiot ovat jopa riippuvaisia siitä (Davenport & Prusak 2000, 2). Data mahdollistaa yrityksille kestävää kilpailukykyä tarjoamalla tilaisuuden ennakoita markkinoilla tapahtuviin liikkeisiin ja vastaamaan niihin. Yrityksen sisäisiin prosesseihin data tuo valmiudet tehokkuuden ja innovatiivisuuden kehittämiseen. (Salo 2013, 136)

Carlsson (2018) kertoo tutkimuksessaan, että tietoperustaiseen talouteen siirryttäessä organisaatioilta vaaditaan muun muassa parempaa ongelmienratkaisukykyä ja päätöksentekoa. Näihin vastatakseen yritykset tarvitsevat asiantuntevia ja kokeneita ihmisiä. Tiedosta on muodostunut tärkeä kilpailuetu. Siksi myös yksittäisen työntekijän osaaminen on saatava hyödynnettyä tehokkaasti koko organisaation tasoisena tietona ja osaamisena (Puusa, Reijonen, Juutu & Laukkanen 2014, 246).

Muutos liiketoimintaympäristössä vaikuttaa niin kaupankäyntiin, asiakkaisiin, markkinoihin kuin kilpailijoihinkin (Rounds 2018, 19). McNallyn (2018) mukaan se johtaa siihen, että myös organisaatioiden on tehtävä muutoksia. Yritykset voivat joutua muuttamaan organisaationsa rakenteita, strategiaansa ja jopa kulttuuriaan. Basun (2015) mielestä organisaatioiden tulisi kyetä tekemään muutoksia näiden lisäksi myös innovaatioihinsa, teknologisiin kyvykkyyksiinsä ja kykyihinsä omaksua ulkoa tulevat muutokset tehokkaasti. Joka tapauksessa tiedon

lisääntyminen pakottaa organisaatiot kaikilla toimialoilla muuttumaan jollakin tapaa ja ajatte-  
lemaan asioita eri tavalla.

## 1.1 Tutkimuksen tavoitteet, rajaukset ja tutkimuskysymys

Tämä kandidaatintyö käsittelee datan ja data-analytiikan hyödyntämistä erään energia-alan  
yhtiön talousyksiköissä. Dataan ja data-analytiikkaan liittyvät tutkimukset ovat lisääntyneet  
viime vuosina, mutta tietoon ja tiedolla johtamiseen liittyviä tutkimuksia on tehty jo pidem-  
pään. Tämä tutkimus rajataan koskemaan yhtiön datan hyödyntämisen vaikutuksia yksittäisen  
työntekijän suorituskykyyn. Tutkimuksia datan ja Big Datan vaikutuksista koko organisaation  
suorituskykyyn on paljon. Sen lisäksi yksilön suorituskyvystä itsessään on tehty aiemmin paljon  
tutkimuksia. Kuitenkin aiheet yhdistettynä aiempia tutkimuksia on niukasti.

Tutkimuksen tavoitteena on selvittää tutkimuskohteena olevan yhtiön talousyksiköiden datan  
ja analytiikan hyödyntämistä ja niiden vaikutuksia yksittäisen työntekijän suorituskykyyn. Tut-  
kimuksen näkökulma on rajattu näiden talousyksiköiden datan ja analytiikan hyödyntämiseen  
ja yksilön suorituskykyyn, jotta tutkimuksella olisi täsmällinen näkökulma sen laajuuteen näh-  
den. Yksilön suorituskyky on valittu myös siksi, että aiheesta löytyy paljon lähdekirjallisuutta  
ja aiempia tutkimuksia. Tutkimusta rajaa lisäksi kyselytutkimuksen kohderyhmä, jona toimii  
yhtiön pyynnöstä organisaation talousyksiköt.

Tutkimuksessa etsitään hypoteesien kautta vastausta tutkimuskysymykseen, joka on seu-  
raava:

*”Miten datan ja data-analytiikan hyödyntäminen vaikuttaa yksilön suorituskykyyn?”*

## 2 TEOREETTINEN TAUSTA

Työhön liittyvät teoriat koskevat data-analytiikkaa, tietoperustaista päätöksentekoa sekä suorituskykyä. Aihealueet muodostavat työn viitekehyksen, johon empiirinen osio perustuu. Tietoon perustuvan päätöksenteon yhteydessä määritellään datalähtöinen toimintakulttuuri, joka liittyy työssä tutkittavaan datan ja analytiikan hyödyntämiseen oleellisesti.

### 2.1 Data ja data-analytiikka

Datalla tarkoitetaan joukkoa erilaisia kvantitatiivisia ja kvalitatiivisia arvoja, joita varastoidaan tietojärjestelmiin. Pelkkä data ei itsessään kerro paljoakaan, koska se toimii rakennusaineena informaatiolle. (Davenport et al. 2000, 2-5) Analytiikassa dataa hyödynnetään laajasti informaation hankkimiseksi, jota voidaan hyödyntää eteenpäin esimerkiksi päätöksentekoon (Davenport et al. 2007, 26).

Teknologia on kehittyessään tuonut mukanaan tehokkaammat välineet datan luomiseen ja sen pysyvämpään tallentamiseen. Lisäksi se on mahdollistanut datan määrän nopean kasvun. Tällaista nopeasti kasvavaa ja monipuolista datamäärää kutsutaan Big Dataksi. (Salo 2013, 10-11) Big Datalla tarkoitetaan siis nimensä mukaisesti suurta määrää dataa, joka on niin monimutkaista ja laajaa, että perinteiset metodit sen käsittelyyn ovat riittämättömiä (Grable & Lyons 2018, 17). Ilmiönä Big Data tarkoittaa datan määrän kasvamista ja monipuolistumista, painetta havaita oleellinen data ja reagoida nopeasti siitä saatuun informaatioon (Salo 2014, 6).

Big Datan strategiat eivät välttämättä sovi tai ole arvokkaita kaikille yrityksille johtuen organisaation rakenteellisista piirteistä tai yrityksestä itsestään. Myöskään kaikki data ei aina ole niin sanottua laadukasta dataa (Corea 2016, 5). Huonoon dataan perustuvat näkökulmat eivät välttämättä ole luotettavia (Smith & Driscoll 2017). Myös huonolaatuisen ja hyvälaatuisen datan yhdistäminen voi vaikuttaa negatiivisesti lopputulokseen. Lisäksi ylimääräisen datan lisääminen malliin ei välttämättä tuo siihen lisäarvoa. Liian suuresta määrästä dataa on vaikeaa tunnistaa merkitsevät osat. Näiden seikkojen vuoksi suurten datamassojen käsittely vaatii paljon

tarkkuutta. (Corea 2016, 5; Davenport et al. 2000, 3) Onkin tärkeää tunnistaa tiettyä tarkoitusta hyödyttävä data. (Gupta 2016, 2)

Data kuvaa ainoastaan sitä mitä tapahtui, eikä se kerro mitä yrityksen pitäisi tehdä seuraavaksi. Päätöksenteko sisältää dataa, mutta vain raaka-aineen muodossa. (Davenport et al. 2000, 3) Datan arvo ei kasva sitä varastoimalla, joten merkittävää siitä tulee vasta silloin, kun se on jalostettu, analysoitu ja välitetty sitä hyödyntäville henkilöille. (Gupta 2016, 2; Markkula & Syväniemi 2015, 72)

Data-analytiikan käsite tuli suosioon 2000-luvun alussa. Käsite määritellään tietojärjestelmien sovellusten avulla tehtäväksi datasettien analysoimiseksi, jolla saadaan tukea päätöksentekoon. (Runkler 2016, 2) Data-analytiikkaan rinnastuva liiketoiminta-analytiikka voi tarkoittaa eri asioita eri organisaatioille, mutta sillä viitataan päätöksentekoa tukevaan analytiikkaan, aivan kuten data-analytiikallakin (Shmueli, Bruse & Patel 2016, 3).

Datasta pyritään siis analytiikan keinoin löytämään tietynlaisia kaavoja ja malleja, joiden pohjalta organisaatio voi tehdä parempia päätöksiä (Gupta 2016, 1). Organisaatio voi näin perustaa liiketoimintansa, johtamisensa ja päätöksentekonsa tosiasioihin (Davenport et al. 2007, 26). Datan analysointiin on tullut uuden ja paremman teknologian ansioista parempia välineitä ja käytäntöjä, jotka helpottavat löytämään organisaatioiden omistamasta suuresta datamassasta malleja, näkökulmia ja trendejä, joita ei muilla tavoin pystyisi havaitsemaan (Gupta 2016, 1).

Analytiikan tuomista hyödyistä huolimatta sen hyväksikäyttö organisaatioissa on vielä vähäistä. Analytiikan keinojen käyttöönottoa voi hidastaa yksinkertaisesti vain uuden aloittamisen vaikeus organisaation johdon ja henkilöstön puolesta sekä teknologisten kykyjen puutteesta. (Markkula et al. 2015, 79)

Data ja sen aiheuttama murros ovat olleet pinnalla jo pitkään, eikä Guptan (2016, 2) mielestä näytä siltä, että tilanne muuttuisi. Datasta on muodostunut uusi pääoman muoto ja innovatiivinen arvon lähde (Corea, 2016, 5). Davenport et al. (2007, 219) uskovat, että tulevaisuudessa



yhä useammat yritykset valjastavat analytiikan osaksi toimintaansa ja saavat siitä kilpailukeinon. Kuitenkin kyky saada analytiikan keinoista kaikki hyödyt irti vaatii yritysjohtolta hyviä tietojohdantamisen taitoja (Davenport et al. 2007, 26).

## 2.2 Tietoon perustuva päätöksenteko

Davenport (2000, 3) määrittelee tiedon sekoitukseksi eri elementtejä, joiden avulla saadaan aikaiseksi uutta tietoa ja kokemusta. Näitä elementtejä ovat esimerkiksi aiempi kokemus, informaatio ja asiantuntemus. Organisaation tieto voidaan ymmärtää opituksi kasaksi normeja, jaetuiksi ymmärryksiksi sekä harjoitteiksi, jotka yhdistyvät työn ja toimijoiden arvokkaiksi tuloksiksi (Newell, Robertson, Scarbrough & Swan 2009, 6).

Tietoperustaiseen talouteen siirtyminen on johtanut siihen, että tiedon ja asiantuntijuuden johtaminen ovat tulleet merkittäviksi. Tieto on tuottavinta pääomaa ja sitä varten organisaatioon tarvitaan taitavaa henkilöstöä hyödyntämään järjestelmiä ja dataa. (Puusa et al. 2014, 245-246) Jos päätöksiä ei tehdä perustuen tietoon, vaan pelkkään intuitioon, kasvavat epävarmuus ja riskit suuriksi. Jotta päätöksenteko olisi tehokasta, on sen perustuttava tietoon ja informaatioon. Tällöin päätöksenteon laatu kohoaa ja epävarmuus sekä riskit pienenevät. (Laitinen 2003, 24-25)

Farrell, Goh, Kahle, Shackell ja White (2017) tutkivat organisaation johdon tunteisiin perustuva päätöksentekoa. He löysivät yhteyden suoritusperusteiseen palkkioon ja päätöksentekoon. Kun johdon palkka on sidottu jollain tapaa suorituskykyyn, muuttuvat heidän päätöksensä taloudellisesti tuottavammiksi. Tämä ei kuitenkaan johdu siitä, että he jättäisivät tunteensa pois päätöksenteossa, vaan päätökset tehdään rationaalisesti, loogisesti ajatellen ja analytiikkaan perustuen. Suorituskykyyn sidottu palkkaus voi hyödyttää Farrell et al. (2017) mukaan sellaisia yhtiöitä, joissa työntekijän asiantuntemus on välttämätöntä. Tällöin työntekijät käyttävät päätöksissään refleksinomaisesti tietoaan ja vanhoja kokemuksiaan.

Farrell et al. (2017) tutkimus osoittaa myös sen, että tunteet liittyvät joka tapauksessa jokaisen ihmisen päätöksiin. Joten sen sijaan, että organisaatioissa pyrittäisiin poistamaan tunteet, pitäisi tunteista johtuvien päätösten seuraukset tunnistaa ja kehittää tapoja vähentää niiden negatiivisia vaikutuksia.

### 2.2.1 Datalähtöinen toimintakulttuuri

Bussen (2012) mukaan empiiriset tutkimukset osoittavat, että datalähtöisellä toimintakulttuurilla voidaan saavuttaa parempia ja nopeampia päätöksiä, parempi kuva asiakkaiden silmissä, vähemmän virheitä sekä erityisen hyvä kilpailuetu. Friskin (2017) tutkimuksessa osoitetaan, että organisaatioiden on muutettava päätöksenteon kulttuuriaan, jotta data-analytiikan ja Big Datan tehokas käyttö voivat lisätä organisaation suorituskykyä mullistavasti.

Yrityskulttuurilla on suuri rooli kriittisen tiedon ja informaation kulun varmistamisessa ympäri organisaatiota. Perinteisesti tieto on kulkenut organisaatiossa vertikaalisesti johdolta työntekijöille. Kuitenkin nykypäivänä tiedon olisi hyvä liikkua myös horisontaalisesti, mikä voidaan saavuttaa muuttamalla organisaation kulttuuria siihen suuntaan. (Dalkir 2005, 185)

Dalkir (2005, 186) määrittelee tietoperustaisen kulttuurin sellaiseksi, jossa tiedon jakaminen koetaan pikemminkin normiksi, kuin odotukseksi. Tiedon liikkuessa työntekijät työskentelevät yhdessä ja jakavat näkökulmiaan, joiden lisäksi heitä palkitaan siitä. Yrityskulttuuri, joka tunnistaa myös äänettömän tiedon ja kannustaa ihmisiä jakamaan sen, ilmenee ympäristössä, jossa ihmisen tieto luokitellaan arvoksi. (Dalkir 2005, 186, 304)

Lavalle, Lesser, Shockley, Hopkins ja Kruschwitz (2011) kertovat tutkimuksensa tuloksista yhdeksi merkittävimmäksi sen, että organisaatioilla suurin haaste analytiikkaan liittyen koskee analytiikan saattamista osaksi omaa johtamistaan ja yrityskulttuuria. Tutkimuksessa nousi esiin muun muassa myös se, että ihmisille on erityisen vaikeaa muuttaa päätöksiensä perustamista omien kokemustensa sijasta dataan.

Datalähtöisen toimintakulttuurin luominen organisaatioon, jossa sitä ei vielä ole, ei ole helppoa. Busse (2012, 12) listaa keinoja tällaisen kulttuurin luomiseksi yrityksissä. Yrityksen on tunnettava statistiikka. Tilastojen, korrelaatioiden ja syy-seuraus-suhteiden ymmärtäessä datan keräysprosessin suunnittelu, tuloksien mittaaminen ja esittäminen sekä käytäntöjen suunnittelu on helpompaa. Myös mittaamisen on oltava osa prosesseja. Datakulttuurin luominen syntyy siitä, että kyetään mittaamaan asioita tarkasti ja esittämään luotettavaa informaatiota. Lisäksi näitä prosesseja tulee kehittää. Faktoihin perustuva päätöksenteko alkaa sellaisten prosessien rakentamisesta, jotka mittaavat tosiasioita ja luovat systeemejä.

Tehokas päätöksenteko perustuu dataan ja informaatioon ensimmäisenä, sitten kokemukseen ja kolmanneksi intuitioon (Busse 2012, 12; Laitinen 2003, 24). Kuitenkin monet johtajat luottavat päätöksissään mieluummin intuitioon ja kokemukseen. Kun organisaatiolla on saatavilla luotettavaa dataa, helpottaa se päätösten perustumista faktoihin tunteiden sijasta. Tämä saa johtajat käyttämään vähemmän aikaa päätöksentekoon, jolloin aikaa jää enemmän strategian suunnitteluun, henkilöstön kehittämiseen ja oikeiden asioiden johtamiseen. (Busse 2012, 12)

### **2.3 Suorituskyky**

Suorituskyvyn termi voidaan määritellä eri tavoin riippuen sen asiayhteydestä tai sitä käsittelevistä ihmisistä (O'Donnel & Duffy 2005, 7). Laitinen (2003, 21) määrittelee suorituskyvyn kyvyksi saada tuloksia aikaan suhteessa määrättyyn tavoitteeseen. Suorituskyvyn mittaaminen tarkoittaa parasta mahdollista suoritusta sekä kykyä suoritusten aikaansaamiseksi, jossa katse kohdistuu tulevaisuuteen. Kun taas suorituksen mittaaminen määritellään toteutuneeseen suoritukseen, eli menneeseen tapahtumaan. (Ukko, Karhu, Pekkola, Rantanen & Tenhu 2007, 3)

Kokonaisvaltaisen suorituskyvyn mittaamisen yleisin järjestelmä on tasapainotettu mittaristo, eli Balanced Scorecard, jonka ovat kehittäneet Kaplan ja Norton vuonna 1992. Tasapainotettu mittaristo, kuten muutkin kokonaisvaltaisen suorituskyvyn mittarit sisältävät taloudellisia ja

ei-taloudellisia mittareita. (Ikäheimo, Laitinen, Laitinen & Puttonen 2014, 129-130). Balanced Scorecard auttaa organisaation johtoa saamaan tietoa yrityksen suorituskyvystä monelta eri osa-alueelta yhtä aikaa. Eli johto saa nopeasti kattavan yleiskuvan organisaation tilanteesta. (Laitinen 2003, 376)

Monissa yrityksissä tavoitteena on parantaa suorituskykyä ja ylläpitää sitä (O'Donnel & Duffy 2005, 8). Suorituskyvyn kehittämisen tarkoituksena on edistää organisaation kykyä vastata erilaisten sidosryhmien tarpeisiin. Näiden tarpeiden tunnistaessa yritys tietää tavoitteet, joihin sen tulee pyrkiä tuotoksen saamiseksi. Esimerkiksi eri sidosryhmille yhteinen tavoite on yrityksen menestyminen, joka rinnastuu tehokkaaseen toimintaan. Tieto ja sen pohjalta tehdyt päätökset liittyvät tähänkin asiayhteyteen siten, ettei yrityksen johtaminen ole tehokasta ilman informaatiota ja tietoa. (Laitinen 2003, 21)

Suorituskyky voidaan jakaa moneen eri osa-alueeseen ja yksi jaottelutapa niiden välillä on sisäinen ja ulkoinen suorituskyky. Sisäinen suorituskyky sisältää esimerkiksi työn laadun ja tuotavuuden, kustannustehokkuuden, osaamisen, käytettävissä olevan teknologian sekä henkilöstön tyytyväisyyden. Ulkoiseen suorituskykyyn voidaan lukea muun muassa kilpailukyky, asiakastyytyväisyys sekä kannattavuus. (Ukko et al. 2007, 6-7) Tässä kandidaatintutkielmassa keskitytään yrityksen sisäiseen suorituskykyyn, koska tutkimuksen kohteena ovat yhtiön sisäiset talousyksiköt ja yksittäinen työntekijä.

Lindin (2015, 17) mukaan organisaatioissa on kokonaisuuksien sijaan hyvä keskittyä yksilötasoon. Työntekijän osalta se voi vahvistaa heidän vaikutusmahdollisuuksiaan, jonka taas odotetaan johtavan yksilön parempaan luovuuteen ja tehokkuuteen. Tabouli, Habtoor ja Nashief (2016) kokoavatkin tutkimuksessaan työntekijän suorituskyvyn tehokkuudesta, työn suunnittelusta, työpanoksesta, luovuudesta ja innovaatiosta. Tutkimuksessa tutkittiin henkilöstöjohtamisen vaikutusta työntekijän suorituskykyyn, jolle löydettiin positiivinen yhteys.

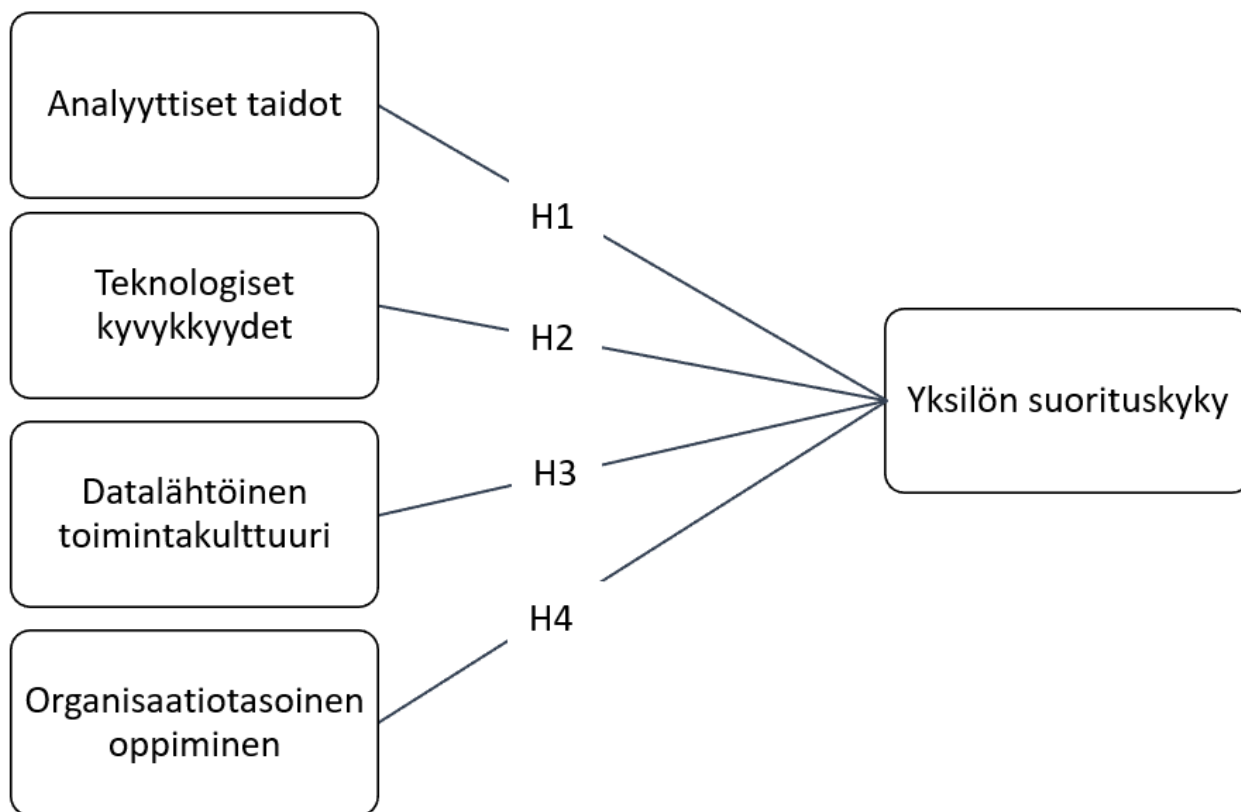
Työntekijän ja tiimien motivoiminen lisää heidän suorituskykyään (O'Donnel & Duffy 2005, 8). Siksi esimiehellä onkin suuri rooli työntekijän suorituskyvyn parantamisessa kannustaen, motivoiden sekä rehellistä palautetta antaen (Sparrow 2012, 132). Myös työntekijän

palkitsemiseminen vaikuttaa suorituskyyyn. Ihmiset panostavat työssään selkeästi niihin asioihin, joiden hyvästä suoriutumisesta palkitaan (Laitinen 2003, 58). Lisäksi työntekijän sitoutumisella organisaatioon on suuri vaikutus hänen suorituskyyynsä (Sparrow 2012, 9-11). Henkilöt, jotka ovat sitoutuneet työhönsä, kokevat henkilökohtaista vastuuta oman työnsä suorituskyyyn ylläpitämisestä (Cooper & Nelson 2007, 144).

Näiden syiden vuoksi motivoimisen ja sitouttamisen, eli henkilöstöjohtamisen tasoisen suorituskyyyn mittaaminen on tärkeää. Se kattaa työntekijän ja tiimin tavoitteiden ja suorituskyyyn mittarien määrittelyn ja tavoitteiden seuraamista. Henkilötason suorituskyyyn mittaamisella voidaan parhaimmassa tapauksessa parantaa henkilöstön johtamista ja työelämän laatua. Nämä positiiviset vaikutukset johtuvat siitä, että työntekijät kokevat voivansa vaikuttaa omiin tavoitteisiinsa ja niiden mittaamiseen. (Ukko et al. 2007, 19)

### **3 TUTKIMUSMALLI JA HYPOTEESIT**

Tässä kappaleessa esitellään tutkimuksen tutkimusmalli sekä siihen pohjautuvat hypoteesit. Tutkimusmalli on esitetty kuviossa 1. Teoriassa esitelty datan ja data-analytiikan hyödyntäminen on jaettu Guptan ja Georgen (2016) sekä Akterin, Wamban, Gunasekaran, Dubeyn ja Childen (2016) tutkimusten pohjalta analyttisiin taitoihin, teknologisiin kyvykkyyksiin, datalähtöiseen toimintakulttuuriin sekä organisaatiotasoiseen oppimiseen. Gupta et al. (2016) todistivat tutkimuksessaan datan ja analytiikan vaikuttavan merkittävästi organisaation suorituskyyyn. Tässä kandidaatintutkielmassa tarkastellaan näiden tekijöiden vaikutusta yksittäisen työntekijän suorituskyyyn, koska koko organisaation suorituskyyyn voidaan ajatella kumuloituvan yksittäisten työntekijöiden suorituskyyvystä.



**Kuvio 1:** Tutkimusmalli

Organisaatio tarvitsee analyttisiä taitoja omaavia henkilöitä hallitsemaan ja analysoimaan dataa (Smith et al. 2017). Tiedon muodostumisella ja hallinnalla on Sujathan ja Krishnavenin (2018) tutkimuksen mukaan positiivinen yhteys työntekijän suorituskkykyyn. Koska analytiikkaa hyödyntämällä yritys voi nopeuttaa ja helpottaa päätöksentekoaan ja johtamistaan (Davenport et al. 2007, 26), kumuloituu se myös yksittäisen työntekijän suoritukseen positiivisesti. Wong ja Sixl-Daniell (2017) osoittivatkin tutkimuksessaan analyttisten taitojen vaikuttavan suuresti yksittäisen työntekijän suoritukseen. Analyttisesti taitava henkilö voi myös jakaa tietoaan ja osaamistaan muun henkilöstön kanssa auttaen näin heitä työtehtävissään. Reychav ja Weisberg (2009) löysivät tutkimuksessaan epäsuoran yhteyden tiedon jakamisen ja yksilön suorituskkyvyn välille. Yhteys ei ollut suora, koska henkilöstön jakama tieto oli hiljaista tietoa. Virtainlahti (2009, 54) määrittelee hiljaisen tiedon teknologisen ulottuvuuden ammatillaisen omaksi tietotaidoksi, jonka hän on saavuttanut kokemuksiensa kautta. Reychavin et al. (2009) tutkimuksessa osoitettiinkin suora yhteys hiljaisen tiedon ja yksilön suorituskkyvyn

välille ja sille, että työntekijät oppivat osaamista toisiltaan ja saavat näin lisättyä suorituskyykyään. Nämä tiedot tukevat ensimmäistä hypoteesia, joka on seuraava:

**H1:** Organisaatiossa esiintyvät paremmat analyyttiset taidot johtavat yksilön parempaan suorituskyykyyn

Teknologiset kyvykkyydet viittaavat Akterin et al. (2016) tutkimuksessa analyytiikan järjestelmien joustavuuteen, yhteensopivuuteen ja modulaarisuuteen sekä siihen, miten helposti järjestelmiä käyttävät saavat niistä tietoa ja tukea yrityksen resursseihin. Teknologiset kyvykkyydet kattavat myös analyyttisen näkemyksen jakamiseen organisaation sisällä. Yhteys teknologisten kyvykkyyksien ja koko yrityksen suorituskyykyille on löytynyt muun muassa Chaen, Kohn ja Prybutokn (2014) tutkimuksessa. Yksittäisen työntekijän suorituskyykyyn liittyen teknologisten kyvykkyyksien vaikutuksen ovat löytäneet Aralin, Brynjolfssonin ja Marshallin (2012), jotka tutkivat tietotekniikan mahdollistaman tiedon lisääntymistä sekä Xiaojun (2017) tutkimuksellaan liittyen tiedonhallintaan perustuviin järjestelmiin. Näitä mukailen toiseksi hypoteesiksi muodostui:

**H2:** Paremmat teknologiset kyvykkyydet johtavat yksilön parempaan suorituskyykyyn

Groverin, Chiangin, Liangin ja Zhangin (2018) tutkimuksen mukaan analyytiikka vaatii toteutukseen vahvaa datalähtöistä toimintakulttuuria ja vahvoja rakenteita datan hallinnoimiseksi. Pitkälle kehittynyt analyytiikka jo itsessään tukee päätöksenteon kehittymistä organisaatiossa ja tuo sinne uusia ja arvokkaita oivalluksia (McGuire, Manyika & Chui 2012). Friskin (2017) mukaan organisaation kulttuurilla on suuri rooli analyytiikan tehokkaassa hyödyntämisessä, ja tutkimuksen mukaan näin organisaation suorituskyykyä voidaankin lisätä merkittävästi. Bussen (2012) tutkimuksessa todettiin datalähtöisen toimintakulttuurin muun muassa vähentävän virheitä, joka taas vaikuttaa jokaisen suoriutumiseen. Lisäksi Lam, Chen & Schaubroeck (2002) toteavat tutkimuksessaan, että työntekijän suorituskyyky nousee, kun hän kokee voivansa osallistua päätöksentekoon. Kolmanneksi hypoteesiksi muotoutui:

**H3:** Datalähtöisellä toimintakulttuurilla on positiivinen vaikutus yksilön suorituskyykyyn

Organisaatiotasoinen oppiminen rakentuu henkilötasoisesta oppimisesta suoraan ja epäsuorasti organisaation muun henkilöstön kanssa. Työntekijän taitojen, kykyjen ja tietojen kehittyessä voi yksilö toimia tuottavammin työssään. Yksilö, joka kokee enemmän oppimista työssään, saa valjastettua oppimisensa suorituskykynsä kasvattamiseksi (Cooper et al. 2007, 78). Jiménez-Jiménez ja Sanz-Valle (2011) löysivät tutkimuksessaan positiivisen yhteyden yrityksen suorituskyvylle ja organisaatiotasoiselle oppimiselle. Näiden syiden vuoksi neljänneksi hypoteesiksi on asetettu seuraava:

**H4:** Organisaatiotasoinen oppiminen vaikuttaa positiivisesti yksilön suorituskykyyn

## 4 TUTKIMUSMENETELMÄ JA -AINEISTO

Tutkimus toteutettiin kvantitatiivisena tutkimuksena lineaarista regressioanalyysia hyödyntäen. Aineisto kerättiin anonymisti sähköisellä kyselylomakkeella, joka lähetettiin tutkimuskohteena olevan yhtiön talousyksiköille. Seuraavaksi esitellään tutkimuksessa käytetty tutkimusmenetelmä, jonka jälkeen tarkastellaan aineiston keruuseen hyödynnetyn kyselylomakkeen muodostamista.

### 4.1 Analyysimenetelmä

Kvantitatiivisella, eli määrällisellä tutkimuksella etsitään muuttujien välisistä suhteista eroja (Vilka 2007, 13). Tällä tutkimustavalla pyritään yleistettävyyteen, syy-seuraus-suhteen selvittämiseen sekä ennustamiseen (Hirsjärvi & Hurme 2001, 22). Määrällinen tutkimus käsittelee numeroita ja käyttäytymistä (Hirsjärvi et al. 2007, 22; Hirsjärvi, Remes & Sajavaara 2013, 137). Tarkoituksena on löytää muuttujien välisiä yhteyksiä, mutta se ei pelkästään riitä. Kvantitatiivisessa tutkimuksessa on pyrittävä myös vastaamaan esimerkiksi siihen, kuinka paljon tätä yhteyttä on. (Vilka 2007, 23)



Tässä tutkielmassa käytetty analyysimenetelmä on lineaarinen regressioanalyysi, joka on yksi vanhimpia ja eniten käytettyjä monimuuttujamenetelmiä (Metsämuuronen 2009, 709). Regressioanalyysillä arvioidaan riippuvuuksia eri muuttujien välillä. Tarkoitus on siis tutkia yhteyttä selitettävän ja selittävien muuttujien kesken sekä mitata niiden riippuvuuksien merkittävyyttä. (Gupta 2016, 50-51; Runkler 2016, 67) Regressioanalyysin selityssaste kuvaa sitä, miten paljon selittävät muuttujat kertovat selitettävän muuttujan vaihtelusta. Selityssaste muodostuu muuttujien välisen yhteyden, eli korrelaatiokertoimen neliöstä (Metsämuuronen 2009, 709).

Lineaarisen regression yleisin malli, eli usean muuttujan analyysi esitetään seuraavasti:

$$Y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_K x_K + \varepsilon \quad (1)$$

Yhtälössä Y kuvaa selitettävää muuttujaa,  $x_2$ - $x_K$  selittäviä muuttujia, joita on K-kappaletta,  $\beta_1$  vakiota ja  $\beta_2$ - $\beta_K$  tuntemattomia parametreja, joita on myös K-kappaletta. Lisäksi  $\varepsilon$  kuvaa jäännöstermiä, eli mallin selittämättä jäämää osaa. Malli ei siis selitä ilmiötä täydellisesti. (Hill, Griffiths & Lim 172, 2012; Metsämuuronen 2009, 721)

Usean muuttujan regressioanalyysiin liittyy taustaoletuksia, joiden tulisi toteutua, jotta malli ei olisi harhainen ja keskivirheet olisivat luotettavia. Näistä ensimmäinen on selitettävän ja selittävien muuttujien välinen lineaarinen yhteys. Toinen oletus liittyy siihen, että jäännöstermin odotusarvon oletetaan olevan nolla, jolloin virheitä ei keskimäärin olisi yhtään. Kolmas oletus koskee jäännöstermin varianssia, jonka odotetaan olevan vakio. Tämä tarkoittaa sitä, että mallin kuuluisi olla homoskedastinen. Neljännes taustaedellytys liittyy jäännöstermien yhteisvaihteluun, eli jäännöstermiparien kovarianssien oletetaan olevan nolla. Viidenneksi selittävät muuttujat eivät saisi olla satunnaismuuttujia eivätkä lineaarisessa yhteydessä toisiinsa. Selittävät muuttujat eivät siis saa korreloida liian vahvasti keskenään. Viimeinen taustaoletus kuvaa sitä, että selitettävän muuttujan ollessa normaalijakautunut, on jäännöstermien noudatettava myös normaalijakaumaa. (Hill et al. 2012, 172-173)

Ennen regressioanalyysin suorittamista käytetään avuksi faktorianalyysia, jonka tarkoituksena on toimia välineenä löytämään muuttujien joukosta keskenään voimakkaasti korreloivat muuttujat. Näistä muuttujista voidaan koota kokonaisuus, jota kutsutaan faktoriksi. Faktorianalyysi voidaan tehdä eksploraatiivisena (EFA, Explorative Factor Analysis) tai konfirmatorisena (CFA, Confirmatory Factor Analysis). Myös pääkomponenttianalyysi (PCA, Principal Component Analysis) kuuluu faktorianalyysien joukkoon. Eksploraatiivisen faktorianalyysin sekä pääkomponenttianalyysin tehtävät ovat samanlaiset, eli etsitään muuttujien joukosta keskinäisen vaihtelun perusteella ryhmiä. Konfirmatorisessa faktorianalyysissa taas haetaan selittävää mallia muuttujien joukosta ja tutkitaan vastaako se valmista mallia. (Hair, Anderson, Tatham & Black 1998, 90-91; Metsämuuronen 2009, 649)

Tässä tutkimuksessa hyödynnetään eksploraatiivista faktorianalyysia, koska malli toimii erityisen hyvin silloin, kun tutkimustilanteessa on valmis teoria, jonka pohjalta voidaan olettaa tutkittavien muuttujien yhdistyvän (Metsämuuronen 2009, 666). Tässä tutkimuksessa onkin tarkoitus muodostaa summamuuttujat tutkimusmallin mukaisesti, joten faktorianalyysilla varmistetaan, että yksittäiset väittämät sopivat niille määriteltyihin kokonaisuuksiin.

## **4.2 Aineiston keräys ja kyselylomakkeen muodostaminen**

Tutkimuksen aineiston keräykseen käytettiin kvantitatiiviselle tutkimukselle tyypillistä kyselyä (Vilkkä 2007, 17), joka toteutettiin kohderyhmälle sähköisesti. Kohderyhmäksi valikoitui tutkimuskohteena olevan yhtiön pyynnöstä organisaation talousyksiköt, joten tutkimus lähetettiin 201 vastaanottajalle. Väittämät olivat kaikille vastaajille tismalleen samat ja samassa järjestyksessä, jotta kysely olisi standardoitu (Hirsjärvi et al. 2013, 193). Koska tutkimuksessa oli tarkoitus selvittää työntekijöiden mielipiteitä sekä käyttäytymistä ja kohderyhmä oli laaja, sopi kysely aineiston keräykseen hyvin sen tehokkuuden ansiosta (Hirsjärvi et al. 2013, 195; Vilkkä 2007, 29).

Kyselylomake muodostettiin aiempien saman tyyppisten tutkimusten perusteella. Lomake on kokonaisuudessaan nähtävissä liitteessä 1. Kaikki väittämät olivat viisi portaisia Likert-

asteikkoja, joilla tutkittiin yksiköiden työntekijöiden mielipiteitä asteikolla 1-5. Likertin asteikkoa käytetään paljon mielipideväittämissä, jolloin vastaaja voi itse arvioida omia käsityksiään kysymyksestä (Metsämuuronen 2009, 70; Vilka 2007, 46). Pienimmällä numerolla vastaaja on täysin eri mieltä ja suurimmalla täysin samaa mieltä. Asteikon keskimäinen arvo kuvaa neutraalia mielipidettä, johon vastaaja voi vastata, jos hän ei esimerkiksi tiedä asiasta.

Lomakkeen väittämät muodostettiin kahdeksasta eri väittämien sarjasta, joista puolet liittyvät yksilön suorituskykyyn ja toiset puolet datan ja analytiikan hyödyntämiseen. Näistä väittämien sarjoista muodostettiin tutkimusmallin pohjalta tutkimuksen eri alueita mittaavat summa-  
muuttujat. Kysymysten kokoamiseen käytettiin aiempia samoja ilmiöitä tutkivia tutkimuksia, jotta mittarit olisivat valideja. Sisällön validiteetti tarkoittaa sitä, että tutkimuksen mittarit ja käsitteet ovat teorian mukaisia ja käsittelevät ilmiötä tarpeeksi laajasti. Mittarien luotettavuudella on suora vaikutus koko tutkimuksen luotettavuuteen. (Metsämuuronen 2009, 74, 126)

Työntekijän suorituskykyä kuvaavat väittämät valittiin Taboulin et al. (2016) tutkimuksen pohjalta. Väittämät koostuvat työn tehokkuudesta, työn suunnittelusta, luovuudesta ja innovatiivisuudesta sekä työhön panostamisesta. Yksilön suorituskyky sopii mittariksi siksi, koska koko organisaation suorituskyvyn voidaan ajatella kumuloituvan yksittäisen työntekijän suorituksesta.

Datan ja analytiikan hyödyntämiseen liittyvät väittämät muodostettiin Guptan et al. (2016) tutkimuksen pohjalta, jossa tutkittiin Big Data analytiikan kyvykkyysien vaikutusta organisaation suorituskykyyn. Valitut väittämät muutettiin koskemaan Big Datan sijasta dataa ylipääntänsä. Ensimmäinen väittämien sarja koostui analyttisista taidoista, joka kuvaa organisaation data-analytiikkaa harjoittavan henkilöstön taitoja. Toinen sarja muodostui datalähtöisestä toimintakulttuurista, joka mittaa yrityksen päätöksenteon perustumista dataan. Teknologisia kyvykkyksiä mittaava väittämien sarja muodostettiin yhdistettynä Guptan et al. (2016) sekä Akterin et al. (2016) tutkimuksista. Sillä kuvataan muun muassa analytiikan järjestelmien yhteensopivuutta ja niistä saatavan tiedon jakamista. Neljäs ja viimeinen väittämien sarja liittyi organisaatiotasoiseen oppimiseen.

## 5 TUTKIMUSTULOKSET

Tutkimuksen tulokset analysoitiin lineaarisella regressioanalyysillä, jonka rinnalla käytettiin faktorianalyysia apuna. Tilastollisiin testeihin hyödynnettiin SAS Enterprise Guide 7.13 ohjelmistoa. Tässä kappaleessa kuvaillaan ensin aineistoa ja muodostetaan selittävien muuttujien faktorit. Viimeisenä esitellään lineaarisen regressioanalyysin tulokset.

### 5.1 Aineiston kuvailu

Tutkimuksen aineisto kerättiin tutkimuskohteena olevan yhtiön talousyksioistä. Aineistonkeruun menetelmänä käytettiin sähköistä kyselylomaketta, joka lähetettiin 201 vastaanottajalle. Kyselyyn oli aikaa vastata 2 viikkoa. Kuitenkaan vielä 1,5 viikon päästä ei vastauksia ollut kertynyt kuin 28 kappaletta. Tämä johtui todennäköisesti syyslomakaudesta, joka sattui juuri samalle aikavälille. Muistutusviestin jälkeen vastauksia kertyi vielä roimasti lisää, joten lopulliseksi vastaajamääräksi vahvistui 46 kappaletta. Vastausprosentiksi muodostui siis 22,9 %. Vastaajien määrään on voinut vaikuttaa myös kyselyn englanninkielinen toteutus. Koska Englanti ei oletettavasti ollut monenkaan vastaajan äidinkieli, on siitä voinut muodostua kynnys kyselyyn vastaamiseksi.

Puuttuvien vastausten osalta aineistoa ei täytynyt seuloa, koska vastaajille kaikki väittämät olivat pakollisia. Aineistosta ei myöskään löytynyt selkeitä outlier-havaintoja, joita olisi täytyntä poistaa. Täten tutkimuksen aineisto muodostui 46 havainnosta, eli kaikista vastaajista. Yksittäisiä väittämiä ei myöskään seulottu vielä tässä vaiheessa, koska faktorianalyysissa tarkasteltiin niiden latautumista faktoreihin ja suoritettiin seulonta sen mukaisesti.

Tutkimuksen analyysimenetelmässä, eli regressioanalyysissa hyödynnettiin summamuuttujia, jotka muodostettiin tutkimusmallin mukaisesti yksittäisistä väittämistä. Mallin selittävät summamuuttujat muodostettiin vasta faktorianalyysin jälkeen, koska faktorianalyysin avulla tarkastettiin muodostuvatko väittämät oletettuihin kokonaisuuksiin. Kuitenkin selitettävää muuttujaa kuvaava summamuuttuja voitiin muodostaa jo aiemmin.

Myöhemmin regressioanalyysissä hyödynnettävä selitettävä muuttuja, yksilön suorituskyky, muodostettiin summamuuttujana yksittäisistä väittämistä. Summamuuttujan muodostamisessa käytettiin keskiarvoa, jotta muuttujasta saatiin jatkuva. Ennen summamuuttujan muodostamista tarkasteltiin väittämien välisiä korrelaatioita Spearmanin järjestyskorrelaatiokerroimen avulla, koska väittämät ovat Likert-asteikollisia ja tulkitaan näin järjestysasteikollisiksi muuttujiksi. Korrelaatioista läheskään kaikki eivät olleet tilastollisesti merkitseviä, ja osa korrelaatioista jäi erittäin pieniksi. Kuitenkin summamuuttujan reliabiliteettia kuvaava Cronbachin alpha muodostui korkeaksi, jopa 0,82. Kuitenkaan väittämät ME2 ja ME3 eivät korreloineet vahvasti minkään toisen muuttujan kanssa. Lisäksi väittämien välisistä korrelaatioista tilastollisesti merkitseviä olivat vain muutama. Täten väittämät päädyttiin poistamaan tarkastelusta sekä summamuuttujasta. Näin summamuuttujan Cronbachin alphaksi muodostui 0,84, joka on hieman suurempi, kuin kyseisten väittämien kanssa.

Summamuuttujan (PERF) perustiedot näkyvät taulukossa 1. Koska muuttujan keskihajonta on 0,4355, sen arvot vaihtelevat todella lähellä keskiarvoa, joka on 3,972. Vaikka muuttuja on muodostettu summamuuttujana väittämien keskiarvona, on sen pienin arvo 3. Vastaaajat ovat siis vastanneet kyselyn suorituskykyyn liittyviin väittämiin melko yksimielisesti suuria arvoja, eikä täysin eri mieltä -vastauksia joukossa esiinny paljoakaan.

**Taulukko 1:** Yksilön suorituskykyä kuvaavan summamuuttujan (PERF) perustiedot

Muuttuja	Keskiarvo	Keskihajonta	Minimi	Maksimi	N
PERF	3,9720	0,4355	3,00	4,7143	46

Yksilön suorituskykyä kuvaavan muuttujan (PERF) jakauma on nähtävissä liitteessä 2. Muuttujan jakaumasta voidaan silmämääräisesti jo huomata sen normaalijakautuminen. Tämä kuitenkin varmistettiin Kolmogorov-Smirnovin testillä. Koska testin p-arvoksi muodostui  $> 0,15$ , jää nollahypoteesi normaalijakaumasta voimaan viiden prosentin riskitasolla, ja voidaan muuttuja todeta normaalijakautuneeksi.

Tutkimusmallin selittävät muuttujat koostuvat yksittäisistä väittämistä, joiden pääkäsitteitä ovat analyyttiset taidot (TS), datalähtöinen toimintakulttuuri (DD), teknologiset kyvykkyydet (TC) ja organisaatiotasoinen oppiminen (OL). Yksittäisten väittämien muuttujia kuvaavat perustiedot ovat esillä liitteessä 3. Väittämien keskiarvot liikkuvat kolmen, eli neutraalin mieltä yläpuolella. Kuitenkin väittämässä DD1, DD2 ja OL1 keskiarvo lähestyvät neljää. Lisäksi väittämässä TC4 keskiarvo putoaa alle kolmen. Väittämien keskihajonnat ovat lähellä yhtä, mutta eivät ylitä sitä. Väittämien arvot eivät siis vaihtelee kovinkaan suuresti keskiarvon ympärillä. Läheskään kaikkiin väittämiin ei ole vastattu pienintä arvoa, eli yhtä. Aineistossa jokaisen väittämän osalta ei siis ole täysin ei mieltä -vastauksia. Kuitenkin täysin samaa mieltä -vastauksia löytyy jokaisen muuttujan osalta.

Koska väittämässä DD1 ja DD2 keskiarvo on miltei neljä, ovat vastaajat ajatelleet keskimäärin melko positiivisesti väittämien käsittelemistä aiheista. Väittämän DD1 saadessa suuria arvoja mieltävät vastaajat organisaation datan yhtiön voimavaraksi. DD2 osalta korkeat arvot kuvaavat sitä, että organisaation päätökset perustuvat mieluummin dataan kuin vaistoon. Datalähtöistä toimintakulttuuria kuvaavan kokonaisuuden muistakin väittämistä on ajateltu positiivisesti. Tämä tarkoittaa sitä, että vastaajat uskovat organisaation toiminnan olevan datalähtöistä.

Väittämän TC4 keskiarvon ollessa alle kolmen, vastaajat ovat mieltäneet, ettei organisaatiossa jaeta analytiikkaan perustuvaa informaatiota yhtenäisesti. Vastaajat olivat samaa mieltä muistakin analyyttisen näkemyksen jakamiseen liittyvistä väittämistä. Voidaan siis olettaa, että vaikka organisaatiossa keskitytään dataan ja sen analysoimiseen, ei tuloksia ja näkemyksiä jaeta kaikkien tietoon.

Analyyttisiin taitoihin (TS) ja organisaatiotasoiseen oppimiseen (OL) liittyen vastaukset ovat jakautuneet melko tasaisesti. Erityisesti kuitenkin organisaatiotasoinen oppimisen osalta ei pienintä arvoa valittu mistään väittämistä. Tämä kuvaa sitä, että vastaajat kokevat uuden tiedon tavoittamisen ja hyödyntämisen organisaatiossaan positiivisena. Analyyttisten taitojen osalta vastaajat ovat luottaneet vahvasti organisaation analytiikkaa harjoittavan henkilöstön

kykyihin. Organisaatiossa myös tarjotaan vastaajien mielestä hyvin koulutusta liittyen dataan ja analytiikkaan.

## 5.2 Faktorianalyysi

Eksploratiivista faktorianalyysia (EFA) käytettiin työkaluna varmistamaan muodostuvatko faktorit tutkimusmallin mukaisiksi kokonaisuuksiksi. Ennen analyysin suorittamista oletettiin, että tiettyyn ryhmään kuuluvat väittämät muodostuvat yhdeksi faktoriksi. Koska tutkimusmallin selittävät muuttujat koottiin kahdesta eri lähteestä, oli faktorianalyysin suorittaminen erityisen tärkeää. Eksploratiivisen faktorianalyysin yksi taustaedellytyksistä on muuttujien välinen korrelaatio, joka tulisi olla yli 0,3 edes joidenkin muuttujien välillä (Metsämuuronen 2009, 667). Tässä tapauksessa jokainen väittämä ylitti tämän rajan, joten taustaedellytys täyttyi. Ainoastaan yksi väittämä DD3 korreloi muita huonommin toisten väittämien kanssa. Väittämää ei kuitenkaan vielä poistettu, koska se todennäköisesti tippuisi pois faktorianalyysissa.

Yleisen nyrkkisäännön mukaan havaintoja tulisi olla viisi kertaa niin paljon kuin muuttujia, mikä liittyy faktorianalyysin taustaoletukseen otoskoon suuruudesta (Hair et al. 1998, 98-99). Tutkimuksen havaintojen määrä, 46, ei siis täytä tätä nyrkkisääntöä. Kuitenkin Metsämuuronen (2009, 666) mukaan otoskoko voi olla pienempi, kunhan muuttujien väliset korrelaatiot ovat korkeita. Havaintojen määrä kaikkein pienimmillään voi olla 50 kappaletta (Hair et al. 1998, 98), johon tämän työn otoskoko miltei ylittää. Guadagnolin ja Velicerin (1988) tutkimuksen mukaan pieni otoskoko ei ole ongelma, kunhan muuttujat ovat latautuneet faktoreihin vahvasti. Pienellä otoskoolla tuloksia on siis tulkittava harkiten niiden luotettavuuden vuoksi (Hair et al. 1998, 99).

Eksploratiivisen faktorianalyysin suorittamiseksi muuttujien tulisi olla myös normaalijakautuneita (Metsämuuronen 2009, 667). Kolmogorov-Smirnovin testin perusteella hylätään jokaisen yksittäisen väittämän kohdalla nollahypoteesi normaalijakaumasta. Yksittäiset väittämät eivät siis ole normaalijakautuneita, johon vaikuttaa erityisesti se, että havaintoja on melko vähän. Visuaalisesti tarkastellen aineisto on kuitenkin kohtalaisen normaalijakautunut.

Multikollinearisuuden esiintyminen aineistossa on ongelma eksploratiivisen faktorianalyysin suorittamiseksi, joka tarkoittaa useiden muuttujien välistä yhteisvaihtelua (Metsämuuronen, 2009, 667). Hairin et al. (1998, 99) mukaan pieni multikollinearisuus olisi kuitenkin toivottavaa, koska faktorianalyysin tarkoituksena on löytää joukko, jossa muuttujat korreloivat keskenään. Väittämien korrelaatiot omissa oletetuissa ryhmissään olivat korkeita. Erityisen korkeita korrelaatioita ei esiintynyt muiden väittämien kanssa. TS ryhmään liittyvät väittämät korreloivat hieman korkeammin DD ja OL väittämien kanssa, mutta korrelaatiot jäivät kuitenkin välille 0,50 – 0,60, joten multikollinearisuus ei muodostunut ongelmaksi.

Faktorointimenetelmäksi valikoitui Unweighted least squares, joka sopii hyvin tilanteeseen, jossa aineistoa on vähän (Metsämuuronen 2009, 673; Zygmunt & Smith 2014, 45). Rotaatiomenetelmänä analyysissa käytettiin Oblique promaxia, joka tarkoittaa vinokulmaista rotaatiota. Vinokulmaisessa rotaatiossa faktoreiden sallitaan korreloivan keskenään (Metsämuuronen 1009, 654). Hairin et al. (1998, 110-111) mukaan menetelmä on hyvä silloin, kun tavoitteena on saada aikaan teoreettisesti merkityksellisiä faktoreita. Todellisuudessa ei ole syytä olettaa, etteivätkö tutkittavat ilmiöt voisi riippua toisistaan, ja useimmiten faktorit joka tapauksessa korreloivatkin edes vähän keskenään. Lisäksi regressioanalyysissä selittävien muuttujien tulisi korreloida keskenään jonkin verran, jotta malli voisi olla hyvä. (Metsämuuronen 2009, 654, 713, 721; Hair et al. 1998, 111)

Eksploratiivisen faktorianalyysin rotatoidut faktorit ja yhteenveto tuloksista ovat nähtävillä taulukossa 2. Kommunaliteetit mittaavat sitä, miten paljon muuttujien varianssista voidaan selittää faktoreilla. Muuttujat latautuvat sitä paremmin faktoreihin, mitä lähempänä niiden kommunaliteetit ovat yhtä. Kommunaliteetti mittaa myös muuttujan luotettavuutta ja siksi sen tulisi saada mahdollisimman suuria arvoja. (Metsämuuronen 2009, 554) Väittämän DD3 kommunaliteetiksi muodostui 0,17, joka on erittäin pieni. Lisäksi muuttujat TS1, TC1, TC3 ja TC5 saivat alle 50 prosentin kommunaliteetit, jotka ovat hieman huonompia. Muuttujia ei kuitenkaan poistettu vielä tarkastelusta, vaan selvitettiin miten ne latautuvat faktoreihin.



**Taulukko 2:** Eksploratiivisen faktorianalyysin tulokset

	<b>OL Faktori 1</b>	<b>DD Faktori 2</b>	<b>TC Faktori 3</b>	<b>TS Faktori 4</b>	<b>Kommunaliteetti</b>
<b>OL1</b>	0,7032				0,5834
<b>OL2</b>	0,8206				0,7724
<b>OL3</b>	0,7376				0,6334
<b>OL4</b>	0,7455				0,6257
<b>OL5</b>	0,6616				0,5862
<b>DD1</b>		0,7704			0,6278
<b>DD2</b>		0,6458			0,5305
<b>DD4</b>		0,6963			0,6131
<b>DD5</b>		0,7245			0,6077
<b>TC1</b>			0,5208		0,3960
<b>TC2</b>			0,6935		0,5997
<b>TC3</b>			0,5579		0,3482
<b>TC4</b>			0,8375		0,7580
<b>TC5</b>			0,6431		0,4580
<b>TC6</b>			0,6054		0,5409
<b>TS1</b>				0,5388	0,3819
<b>TS2</b>				0,6832	0,5839
<b>TS4</b>				0,7897	0,7759
<b>TS5</b>				0,7177	0,8580
<b>Ominaisarvo</b>	6,9059	2,2627	1,7475	1,0377	
<b>Cum. %</b>	0,5777	0,7670	0,9132	1,0000	
<b>Cronbach alpha</b>	0,8891	0,8224	0,8351	0,8351	

Pienen kommunaliteetin omaava väittämä DD3 ei latautunut mihinkään faktoriin kunnolla. Sen alkuperäiseen oletettuun ryhmään lataus oli suurin, eli 0,36. Väittämä pudotettiin kuitenkin tarkastelusta sen pienen kommunaliteetin ja faktorilatauksen vuoksi. Guadagnolin et al. (1998) mukaan muuttujan lataus faktorille ei saisi olla pienempi kuin 0,4, etenään pienellä otoskoolla. Hair et al. (1998, 111) kertovat, että latauksen ehdoton minimi on 0,3, mutta 0,5 on jo merkitseviä. Tämä koskee kuitenkin vain suurta otoskokoja. Noin 50 muuttujan otoskoolla merkittäviä latauksia ovat yli 0,75 menevät arvot.

Väittämä TS3 latautui kahdelle faktorille tasan yhtä paljon. Lataus sen alkuperäiseen oletettuun ryhmään oli 0,55 ja faktorille numero kaksi 0,54. Lataukset ovat suuret, jonka vuoksi

väittämä päätettiin pudottaa tarkastelusta. Väittämän poistaminen oli tärkeää myös siksi, jotta väittämät saataisiin eroteltua kokonaisuuksiin tutkimusmallin mukaisesti.

Väittämän TS3 poistamisen jälkeen faktorianalyysi suoritettiin uudelleen ja näin muuttujat latautuivat selkeästi tietyille faktoreille. Alkuperäisen oletuksen mukaisesti väittämät jakautuivat odotettuihin kokonaisuuksiin. Kuitenkin vain muutaman väittämän osalta faktoriin latautumisen 0,75 raja-arvo ylittyy. Nämä väittämät ovat OL2, OL4, DD1, TC4, ja TS4. Näiden väittämien latausten voidaan siis sanoa olevan varmasti merkittäviä näin pienellä otoskoolla. Vaikka useat väittämät alittavat 0,75 raja-arvon ovat niistä monen kommunaliteetti kuitenkin korkea, joten näiden väittämien voidaan sanoa sopivan tähän kokonaisuuteen siitä huolimatta.

Väittämien TC1, TC3 ja TS1 osalta lataukset jäivät 0,5 vaiheille sekä kommunaliteetit noin 30 prosenttiin. Nämä väittämät voisi poistaa tarkastelusta, mutta niiden ollessa mukana omassa kokonaisuudessaan, saavat uudet summamuuttujat korkeat Cronbachin alphet. Tämä tarkoittaa korkeaa reliabiliteettia jokaiselle kokonaisuudelle, jonka takia väittämät jätetään mukaan tarkasteluun. Nämä väittämät voivat kuitenkin horjuttaa faktorianalyysin validiteettia sekä myöhempien tutkimustulosten luotettavuutta.

Ominaisarvo mittaa faktorin hyvyttä tarkastellen kaikkien muuttujien latauksia faktorille. Ominaisarvon suuruuteen liittyvä nyrkkisääntö on, että sen olisi oltava vähintään yksi. Kuitenkin tällaiset nyrkkisäännöt eivät ole yksiselitteisiä ja voivat erityisesti vaihdella eri faktorointimenetelmän kohdalla. (Metsämuuronen 2009, 669)

Taulukon 2 alareunassa näkyvät faktoreiden ominaisarvot sekä kumulatiiviset prosentit. Kumulatiivinen prosentti liittyy ominaisarvoon ja kertoo paljon faktori selittää muuttujien vaihtelusta. Taulukosta nähdään, että ensimmäinen faktori selittää yli puolet muuttujien vaihtelusta. Kun taas viimeinen selittää sata prosenttisesti muuttujien vaihtelun.

### 5.2.1 Summamuuttujien muodostaminen faktorianalyysin pohjalta

Yksittäiset väittämät muodostuivat faktorianalyysin avulla ennalta oletettuihin kokonaisuuksiin, joten dataan ja analytiikkaan liittyvät summamuuttujat voitiin muodostaa faktoreiden pohjalta. Summamuuttujat muodostettiin väittämien keskiarvona, jotta muuttujista saatiin jatkuvia. Ennen tätä tarkasteltiin kuitenkin väittämien korrelaatioita ryhmissään sekä uuden summamuuttujan reliabiliteettia.

Organisaatiotasaisen oppimisen väittämien ryhmän (OL) kaikki korrelaatiot olivat tilastollisesti merkitseviä viiden prosentin riskitasolla. Kaikki väittämät myös korreloivat keskenään vähintään 0,5 korrelaatiokertoimella. Jopa erittäin korkea, eli yli 0,8 korrelaatio ilmeni kahden väittämän välillä (OL1 ja OL2). Lisäksi summamuuttujan reliabiliteettia kuvaava Cronbachin alphasiksi muodostui 0,89, joka ei paransisi yhtäkään muuttujaa poistamalla.

Datalähtöistä toimintakulttuuria kuvaavan kokonaisuuden (DD) korrelaatiot olivat myös tilastollisesti merkitseviä. Korrelaatiot vaihtelivat välillä 40 – 60, joita sanotaan Metsämuurosen (2009, 371) mukaan kohtuullisiksi korrelaatiokertoimiksi. Korrelaatioita tarkasteltiin faktorianalyysissä poistuneen väittämän DD3 kanssa sekä ilman. Väittämä ei korreloinut vahvasti minkään toisen väittämän kanssa, ja merkitseviä korrelaatioita oli vain yhden toisen väittämän kanssa. Väittämän DD3 kanssa summamuuttujan Cronbachin alpha olisi 0,79, mutta ilman sitä 0,82.

Kolmannen faktorin osoittama kokonaisuus kuvaa teknologisia kyvykkyyksiä (TC). Kaikki väittämien väliset korrelaatiot olivat tilastollisesti merkitseviä. Väittämistä TC5 korreloi heikokimmin väittämien TC1 ja TC3 kanssa. Kuitenkin nämä väittämät korreloivat toisten väittämien kanssa hyvin ja summamuuttujan Cronbachin alphasiksi muodostui 0,84, joka ei kasvaisi yhtään väittämää poistamalla.

Viimeisen kokonaisuuden, analyyttiset taidot (TS), Cronbachin alpha oli 0,84. Tämänkään summamuuttujan reliabiliteetti ei olisi parantunut poistamalla väittämiä tarkastelusta. Kaikki korrelaatiot olivat tilastollisesti merkitseviä sekä korkeita. Väittämän TS1 kanssa muodostuvat

korrelaatiot olivat hieman heikompia, kuin muiden väittämien väliset korrelaatiot. Väittämää ei kuitenkaan päädytty poistamaan kokonaisuudesta, koska korrelaatiot olivat kuitenkin tilastollisesti merkitseviä ja Cronbachin alpha korkea.

Uusien summamuuttujien perustiedot ovat esillä taulukossa 3. Kuten yksittäisten väittämien osalla, uusien summamuuttujien keskihajonta ei ole suurta, joten arvot liikkuvat lähellä keskiarvoja. Jakaumat vaikuttavat perustietojen osalta muutoin hyviltä, koska keskiarvot ovat lähellä kolmea ja suurimmat sekä pienimmät arvot lähellä ääripäitä. Jakaumia tarkasteltiin myös visuaalisesti kuvaajien perusteella, jotka ovat nähtävissä liitteessä 2.

**Taulukko 3:** Summamuuttujien perustiedot

Muuttuja	Keskiarvo	Keskihajonta	Minimi	Maksimi	N
OL	3,6913	0,5715	2,4	5,0	46
DD	3,6196	0,6492	1,5	4,75	46
TC	3,2391	0,6000	2,0	4,5	46
TS	3,2826	0,6425	1,75	5,0	46

Muuttujien normaalijakautumiset testattiin Kolmogorov-Smirnovin testillä. Muuttujan TS kohdalla p-arvoksi muodostui 0,148, joka tarkoittaa nollahypoteesin voimaan jäämistä, eli muuttuja on normaalijakautunut. Muiden muuttujien, DD, TC ja OL, p-arvoiksi muodostuivat järjestyksessä 0,017, < 0,01 ja 0,02. Eli oletus normaalijakaumasta hylätään näiden summamuuttujien osalta. Vaikka testin perusteella muuttujat DD, TC ja OL eivät ole normaalijakautuneita, voidaan silmämääräisesti nähdä muuttujien olevan kuitenkin lähellä normaalijakaumaa. Muuttujamuunnoksille ei siis ollut tarvetta. Aineistossa ei myöskään näkynyt selkeitä outlier-havaintoja, jota olisivat voineet vaikuttaa jakaumiin.

### 5.3 Lineaarinen regressioanalyysi

Lineaarisella regressiolla tutkitaan tutkimusmallin mukaisesti määriteltyjen hypoteesien toteutumista, eli datan ja data-analytiikan hyödyntämisen vaikutusta yksittäisen työntekijän suorituskyykyyn. Selitettävänä muuttujana toimii yksilön suorituskyyky ja selittävinä analyttiset taidot, teknologiset kyvykkyydet, datalähtöinen toimintakulttuuri sekä organisaatiotasoinen oppiminen. Estimointimenetelmänä analyysissä käytettiin pienimmän neliösumman menetelmää (OLS, Ordinary Least Squares), joka minimoi havaintojen ja regressiosuoran välisen etäisyyden neliöt. Menetelmällä on siis tarkoitus löytää suora, jolla tämän etäisyyden neliösumma on niin pieni kuin mahdollista (Hill et al. 2012, 51).

Regressioanalyysin taustaoletusten täyttymistä voidaan tarkastella ennen estimointia ja itse estimoinnin yhteydessä. Ennen analyysin suorittamista tarkasteltiin muuttujien välisiä korrelaatioita. Metsämuurosen (2009, 713) mukaan regressioanalyysia suorittaessa selittävien muuttujien tulisi korreloida kohtuullisesti toistensa kanssa, joka tarkoittaa 0,4 – 0,6 asteista korrelaatiokerrointa. Korrelaatioita tarkasteltiin Pearsonin tulomomenttikorrelaatiokertoimen avulla. Muuttujien väliset korrelaatiot ovat nähtävillä taulukossa 4, josta huomataan korrelaatioiden jääneen melko alhaisiksi. Ainoastaan analyttisiä taitoja kuvaavan muuttujan (TS) ja selitettävän muuttujan (PERF) välinen korrelaatio on kohtalaisella tasolla. Selitettävän muuttujan kanssa huonoiten korreloi teknologiset kyvykkyydet (TC), joka ei ollut edes tilastollisesti merkitsevä, toisin kuin muiden korrelaatioiden kohdalla. Korrelaatiotaulukosta nähdään myös, etteivät selittävät muuttujat korreloi keskenään vahvasti, joten multikollineaarisuus ei vaikuttaisi olevan ongelma.

**Taulukko 4:** Muuttujien väliset korrelaatiot

	PERF	TS	TC	DD
TS	0,50677			
TC	0,25391	0,34198		
DD	0,36996	0,44994	0,34567	
OL	0,37400	0,49405	0,44254	0,32031

Heikot korrelaatiot selitettävän muuttujan kanssa kertovat lineaarisen yhteyden puuttumisesta. Oletus lineaarisuudesta tarkistettiin kuitenkin myös estimoinnin yhteydessä residuaalikuvaajien avulla, jotka ovat nähtävillä liitteessä 4. Residuaalit suhteessa selittäviin muuttujiin ovat kuvaajien perusteella riippumattomia selittävien muuttujien arvoista. Jännöstermien vaihteluväli ei kuitenkaan pysy samana siirryttäessä kuvaajalla, joten muuttujien lineaarinen yhteys ei ole täysin selkeää. Residuaalikuvaajista voidaan nähdä myös outlier-havaintoja, joita ei aiemmista aineiston tarkasteluista huomattu. Outliereita ei ollut vain selkeästi muutamaa, joten niitä ei siksi päädytty kuitenkaan poistamaan. Lisäksi outlier-havaintojen poistaminen olisi rajannut aineiston todella pieneksi. Residuaaleista suhteessa ennustettuun arvoon nähdään, että jännöstermit ovat satunnaisesti levittäytyneet eivätkä vaihtelee systemaattisesti. Kuitenkin oletus homoskedastisuudesta tarkistettiin lisäksi Whiten testillä. P-arvoksi muodostui 0,6861, joten viiden prosentin riskitasolla nollahypoteesi homoskedastisuudesta jäi voimaan.

Estimoinnin yhteydessä havaitut toleranssit kertovat multikollineaarisuudesta, jota on syytä epäillä toleranssien ollessa pieniä (Metsämuuronen 2009, 728). Toleranssit olivat kuitenkin korkeita, joten multikollineaarisuus voitiin poissulkea. Regressioanalyysin viimeisen taustaoletuksen mukaan residuaalienkin tulisi olla normaalijakautuneita, koska selitettävä muuttuja (PERF) todettiin normaalijakautuneeksi aiemmassa kappaleessa 5.1. Residuaalikuvaajista voidaan havaita jännöstermien normaalijakautuminen, mutta asia tarkistettiin vielä Kolmogorov-Smirnovin testillä. Nollahypoteesi normaalijakautumisesta jäi voimaan viiden prosentin riskitasolla, koska testin P-arvoksi muodostui  $> 0,15$ . Usean muuttujan regressioanalyysin taustaedellytykset siis täyttyivät lineaarista yhteyttä lukuun ottamatta.

Lineaarisen regressioanalyysin tulokset ovat tiivistettynä taulukossa 5. Koko mallin merkitsevyyttä tutkitaan F-testin avulla (Hill et al. 2012, 226). Viiden prosentin riskitasolla hylätään nollahypoteesi siitä, että yhdenkään muuttujan kerroin ei poikkeaisi nolasta. Tämä tarkoittaa sitä, että ainakin yksi kertoimista poikkeaa merkittävästi nolasta, joten selittävät muuttujat selittävät yksilön suorituskykyä tilastollisesti merkitsevästi. Mallin selitysaste ( $R^2$ ) kuvaa, miten paljon muuttujat yhdessä selittävät selitettävän muuttujan vaihtelusta (Metsämuuronen 2009, 709). Selitysasteeksi muodostui 29,74 prosenttia, joka jäi melko pieneksi. Tämä on kuitenkin Hairin et al. (1998, 165) mukaan tyypillistä tämän kokoiselle aineistolle tällä muuttujien määrällä. Korjattu selitysosuus ( $R^2_{adj}$ ) huomioi muuttujien määrän sekä otoskoon (Metsämuuronen 2009, 720). Mallin korjattu selitysosuus oli 22,88 prosenttia, joka ei suuresti poikkea mallin selitysasteesta. Muuttujien määrällä ja otoskolla ei siis ollut merkittävän suurta vaikutusta mallin selitysasteelle.

**Taulukko 5:** Lineaarisen regressioanalyysin tulokset

	<b>B</b>	<b>Keskivirhe</b>	<b>T-arvo</b>	<b>Pr &gt;  t </b>
<b>Vakio</b>	2,36040	0,4455	5,30	<0,001
<b>TS</b>	0,24594	0,10949	2,25	0,0301
<b>DD</b>	0,10571	0,10115	1,05	0,3021
<b>TC</b>	0,01040	0,10950	0,09	0,9248
<b>OL</b>	0,10512	0,12195	0,86	0,3937
<b>Sopivuus</b>	<b><math>R^2</math></b>	<b>Korjattu <math>R^2</math></b>	<b>F-arvo</b>	<b>Pr &gt; F</b>
	0,2974	0,2288	4,34	0,0051

Taulukosta 5 nähdään, että vain yksi selittävästä muuttujista, analyttiset taidot (TS), on tilastollisesti merkitsevä viiden prosentin riskitasolla. Tämä selviää T-testin avulla, jolla tutkitaan yksittäisen muuttujan tilastollista merkitsevyyttä (Hill et al. 2012, 185). Muuttuja selittää siis yksilön suorituskykyä tilastollisesti merkitsevästi, jonka lisäksi sillä on suurin vaikutus selitettävään muuttujaan. Toiseksi vahvin vaikutus on datalähtöistä toimintakulttuuria kuvaavalla muuttujalla (DD) ja pienin vaikutus on teknologisia kyvykkyksiä kuvaavalla muuttujalla (TC), jolla myös T-arvo jäi pienimmäksi.

Regressioanalyysin perusteella tutkimuskohteena olevan yhtiön talousyksiköiden datan ja data-analytiikan hyödyntämisellä on tilastollisesti merkitsevä vaikutus yksilön suorituskykyyn. Datan analysoinnilla voidaan luoda organisaatioon tehokkuutta ja uusia innovaatioita, jotka jo itsessään vaikuttavat yksilöiden suoriutumiseen. Koska päätökset voidaan perustaa faktoihin, saadaan päätöksistä laadukkaampia ja niitä voidaan tehdä nopeammin. Näin organisaatiossa voidaan saada aikaan heti oikeanlaisia ratkaisuja, joita ei myöhemmin tarvitse korjailla. Yksittäisten työntekijöiden suoriutuminen siis paranee, koska johdon tekemiä päätöksiä ei tarvitse odotella eikä virheiden korjaamiseen kulu turhaan työaika.

Regressioanalyysin avulla saatiin myös tutkimusmallin ensimmäistä hypoteesia (*H1: organisaatiossa esiintyvät paremmat analyttiset taidot johtavat yksilön parempaan suorituskykyyn*) tukeva vastaus, joka on linjassa aiempien tutkimusten kanssa. Esimerkiksi myös Wong et al. (2017) osoittivat tutkimuksessaan analyttisten taitojen olevan yksi merkittävimmistä tekijöistä yksilön suorituskykyä ennustaessa. Regressioanalyysin tulosta selittää analyttisiä taitoja kuvaavan muuttujan (TS) vahva korrelaatio selitettävän muuttujan (PERF) kanssa, joka nähtiin korrelaatiomatriisista taulukossa 4. Osaavan henkilökunnan ansiosta organisaatio kykenee hyödyntämään keräämäänsä dataa tehokkaasti, joka olisi muutoin hyödytöntä. Koko organisaatio hyötyy siis taitavista yksilöistä, jotka voivat osaamisellaan auttaa myös muuta henkilöstöä suoriutumaan paremmin työstään. Osaamista voidaan jakaa organisaatiossa hiljaisena tietona, jolla voi näin olla merkittäviä vaikutuksia muun henkilöstön suorituskykyyn.

Toiseen hypoteesiin (*H2: Paremmat teknologiset kyvykkyudet johtavat yksilön parempaan suorituskykyyn*) regressioanalyysi ei antanut oletettua vastausta, vaikka näyttöä hypoteesin toteutumiselle oli paljon aiempien tutkimuksien perusteella. Tulokseen vaikuttavat oletettavasti heikommat korrelaatiot summamuuttujan sisällä, vaikka muuttujan reliabiliteetti olikin korkea. Muuttuja ei myöskään ollut täysin normaalijakautunut, ja yksittäiset väittämät saivat hyvin erilaisia arvoja verrattuna muiden kokonaisuuksien väittämiin, mistä kertoivat jo faktorianalyysin heikot kommunaliteetit.

Tutkimusmallin kolmaskaan hypoteesi (*H3: Datalähtöisellä toimintakulttuurilla on positiivinen vaikutus yksilön suorituskykyyn*) ei toteutunut. Datalähtöistä toimintakulttuuria kuvaavalla



muuttujalla (DD) oli kuitenkin eniten vaikutusta selitettävään muuttujaan (PERF) verrattuna muihin muuttujiin, jotka eivät olleet tilastollisesti merkittäviä. Summamuuttujan jakauma oli myös hieman vinoutunut oikealle ja mukana on voinut olla tuloksia vääristäviä outlier-havain- toja. Hypoteeseista viimeinen (*H4: Organisaatiotasoinen oppiminen vaikuttaa positiivisesti yksilön suorituskyykyyn*) ei myöskään toteutunut. Estimoinnin yhteydessä residuaalien sirontaku- vioita tarkastellessa organisaatiotasoinen oppimisen (OL) muuttujan kohdalla outlier-havain- toja oli paljon, jotka ovat voineet vääristää tulosta. Muuttuja OL ei myöskään korreloinut vah- vasti selitettävän muuttujan (PERF) kanssa, joka on myös oletettavasti vaikuttanut regressio- analyysin tulokseen.

Tutkimustulokset eivät ole yksiselitteisiä ja eivätkä välttämättä anna todellista kuvaa asioiden yhteyksistä. Tuloksiin on oletettavasti vaikuttanut usean muuttujan samanaikainen tarkastelu. Jos regressioanalyysi olisi toteutettu yhdellä selittävällä muuttujalla tapaus kerrallaan, olisivat muuttujat DD ja OL olleet tilastollisesti merkitseviä. Metsämuurosen (2009, 635) mukaan mo- nimuuttuja-aineistoja tutkittaessa on havaintoja oltava riittävästi ja jopa alle 200 havainnon tapauksissa on suhtauduttava tuloksiin kriittisesti. Havaintojen määrä on voinut tässä tapauk- sessa vaikuttaa regressioanalyysin tuloksiin, vaikkakin viisi havaintoa jokaista muuttujaa koh- den täytyi. Tulosten paikkansapitävyyteen voivat vaikuttaa myös outlierit, lineaarisen yhtey- den puuttuminen sekä otannan satunnaisuus (Metsämuuronen 2009, 636-637). Koska havain- not kerättiin yhden organisaation talousyksiköistä, on mahdollista, että yksittäiset vastaajat ovat hyvin samankaltaisia ja samaa mieltä asioista, kuin olisi tilanteessa, jossa otanta on täysin satunnainen. Tällainen tilanne johtaa Metsämuurosen (2009, 638) mukaan tapaukseen, jota kutsutaan ryvästymiseksi. Termi tarkoittaa sitä, että joukossa varianssi jää paljon pienem- mäksi, kuin täysin satunnaisessa otoksessa.

#### **5.4 Tutkimustulosten luotettavuus**

Tulosten luotettavuuteen vaikuttavat reliabiliteetti sekä validiteetti (Metsämuuronen 2009, 74). Reliabiliteetti viittaa tutkimuksen toistettavuuteen ja taipumukseen antaa ei-sattuman- varaisia tuloksia. Tutkimuksen voidaan ajatella olevan tarkka ja luotettava silloin, kun

toistettaessa tutkimus antaa samanlaiset tulokset, vaikka tutkija olisikin eri. Vaikka reliabiliteettia voidaan tarkastella tutkimuksen jälkeenkin, on sitä syytä arvioida myös tutkimuksen aikana. (Vilka 2007, 149) Reliabiliteetti voidaan laskea kolmella eri tavalla: toistomittauksella, rinnakkaismittauksella tai mittarin sisäisten yhteyksien kautta (Metsämuuronen 2009, 75).

Toistomittaus vaatisi kyselytutkimuksen suorittamista uudelleen tietyn ajan kuluttua, joka ei tässä tapauksessa olisi mahdollista. Kuitenkin tutkimuksen reliabiliteettia tarkasteltiin summamuuttujien yhteydessä Cronbachin alphan avulla, jolla mitataan mittarin sisäistä yhteyttä (Metsämuuronen 2009, 76). Jokaisen summamuuttujan reliabiliteettikerroin muodostui korkeaksi, joten muuttujat olivat luotettavia. Selitettävän muuttujan, yksilön suorituskyvyn, tapauksessa summamuuttuja muodostettiin todella monen eri väittämän keskiarvona. Usea väittämän lisääminen summamuuttuun johtaa Hairin et al. (1998, 118) mukaan Cronbachin alphan nousemiseen, vaikka korrelaatiot eivät olisikaan vahvoja. Selitettävän summamuuttujan luotettavuuteen onkin siksi syytä suhtautua varauksella.

Validiteetilla tarkoitetaan tutkimusmenetelmän tai mittarin kykyä mitata juuri sitä, mitä oli tarkoitus mitata (Hirsjärvi 2013, 231). Ulkoinen validiteetti tarkoittaa sitä, miten hyvin tutkimus voidaan yleistää. Sisällön validiteetti taas käsittelee sen, että ovatko tutkimuksen käsitteet ja mittarit teorian mukaisia ja käsittelevätkö ne ilmiötä tarpeeksi laajasti. (Metsämuuronen 2009, 74) Kun tutkimuksessa ei esiinny systemaattisia virheitä ja käsitteet eivät johda harhaan, on tutkimuksella hyvä validiteetti (Vilka 2007, 150).

Tutkimuksen tutkimusmalli ja kyselylomakkeen väittämät muodostettiin aiempien tutkimusten perusteella, jotta sisällön validiteetti toteutuisi. Tutkimuksen mittarit vastaavat siis teoriaa. Väittämien validiteettia on voinut horjuttaa se, että vastaajat ovat käsittäneet niiden tarkoituksen toisin kuin tutkija on ne tarkoittanut. Esimerkiksi teknologisia kyvykkyyksiä kuvaavien väittämien kohdalla ovat kysymykset voineet olla liian teknisiä toisille vastaajille. Lisäksi väittämiin vastattiin asteikolla 1–5, jolloin keskimmäiseen vaihtoehtoon, neutraaliin mielipiteeseen, on voitu vastata liian helposti. Kyselytutkimuksen vastausprosentti jäi melko alhaiseksi, joka vaikuttaa siihen, ettei otos edusta kovinkaan hyvin perusjoukkoa.

Tutkimustulosten yleistettävyyttä tutkittavan yrityksen talousyksiköihin on siis harkittava. Lisäksi tutkittavan aineiston pieni koko on voinut vaikuttaa tulosten tarkkuuteen.

Metsämuurosen (2009, 729) mukaan regressioanalyysin tulosten luotettavuuteen vaikuttavat aineistossa esiintyvät outlier-havainnot, jotka vääristävät tuloksia. Aineistossa esiintyi paljon outliereita, joista yksikään ei ollut selkeästi muita suurempi tai selkeämmin poistettavissa. Havaintoja ei päädytty poistamaan aineiston koon vuoksi, joten tutkimustulokset eivät ole täysin luotettavia ja niihin on suhtauduttava varauksella. Aineiston koko vaikuttaa myös faktorianalyysin luotettavuuden tarkasteluun. Faktorianalyysin stabiilisuutta olisi voitu tarkastella jakamalla aineisto kahtia ja suorittamalla faktorianalyysi uudelleen. Näin olisi varmistettu faktoroiden muodostuneen oikein. Aineistoa oli alun perinkin liian vähän faktorianalyysin suorittamiseksi, joten faktorianalyysin suorittaminen vielä pienemmällä otoskolla ei olisi ollut suotavaa. Faktorianalyysin validiteettiin vaikuttavat myös pienet kommunaliteetit ja heikot faktorilataukset omaavat väittämät, jotka jätettiin tarkasteluun mukaan.

## 6 YHTEENVETO JA JOHTOPÄÄTÖKSET

Tässä tutkimuksessa tutkittiin erään energia-alan yhtiön talousyksiköiden datan ja data-analytiikan hyödyntämistä ja sen vaikutusta yksittäisen työntekijän suorituskyykyyn. Tutkimus toteutettiin kyselytutkimuksella, johon saatiin vastauksia 46 kappaletta 22,9 vastusprosentilla. Kyselylomakkeen tuloksia analysoitiin lineaarisen regressioanalyysin avulla, jonka tuloksiin on kuitenkin suhtauduttava kriittisesti pienen otoskoon vuoksi. Tutkimuksen päätutkimuskysymykseen etsittiin vastaus hypoteesien kautta, jotka muodostettiin tutkimusmallista. Tutkimusmalli koottiin aiempien tutkimusten pohjalta, joten datan ja analytiikan hyödyntäminen jaoteltiin neljään osaan, joita ovat analyttiset taidot, teknologiset kyvykkyydet, datalähtöinen toimintakulttuuri sekä organisaatiotasoinen oppiminen.

Tutkimustulokset antoivat vastauksen päätutkimuskysymykseen, joka oli seuraava:

*”Miten datan ja data-analytiikan hyödyntäminen vaikuttaa yksilön suorituskyykyyn?”*

Lineaarisen regressioanalyysin mukaan datan ja data-analytiikan hyödyntäminen vaikuttavat yksilön suorituskyykyyn tilastollisesti merkittävästi, vaikkakin mallin selitysosuus jäi hieman matalaksi. Tutkimuksella osoitettiin siis tutkimuskohteen talousyksiköiden datan hyödyntämisellä olevan positiivinen vaikutus yksittäisen työntekijän suorituskyykyyn.

Tutkimuskysymystä tarkastelevista hypoteeseista vain yksi toteutui. Tulosten mukaan organisaatiossa esiintyvillä analyyttisillä taidoilla on tilastollisesti merkitsevä vaikutus yksittäisen työntekijän suorituskyykyyn. Talousyksiköiden analyyttisiä taitoja omaavalla henkilökunnalla on siis keskeinen rooli organisaation datan analysoinnissa ja valjastamisessa organisaation hyödyksi. Kyselytutkimuksen perusteella talousyksiköissä uskotaankin vahvasti analytiikkaa harjoittavan henkilöstön taitoihin. Osaavien yksilöiden taidot voivat vaikuttaa heidän oman suorituskyykynsä lisäksi organisaation muidenkin työntekijöiden suoriutumiseen. Nykyajan tietoperustaisessa taloudessa onkin erityisen tärkeää saada valjastettua yksilön osaamista koko organisaatiolle. Tutkimuskohteen talousyksiköiden sisällä osaamista saatetaan jakaa hiljaisen tiedon kautta, sillä tutkimuksessa paljastui, ettei organisaatiossa analyyttistä näkemystä jaeta tarpeeksi selkeästi ja julkisesti. Aiheesta voisi suorittaa jatkotutkimuksia, jotta selviäisi miten osaamista ja tietoa jaetaan yhtiön talousyksiköissä. Kyselytutkimuksen perusteella selvisi lisäksi se, että yhtiö tarjoaa dataan ja analytiikkaan liittyvää koulutusta henkilöstölleen, joten kaikki työntekijät voivat oppia analyyttisiä taitoja ja siten parantaa suorituskyykyään.

Vaikka muut hypoteesit eivät toteutuneet, on niillä kuitenkin regressioanalyysin mukaan jonkinlaista vaikutusta yksilön suorituskyykyyn. Erityisesti datalähtöinen toimintakulttuuri vaikuttaa yksilön suorituskyykyyn, vaikkakin ei tilastollisesti merkittävästi. Organisaation datalähtöinen toimintakulttuuri tehostaa ja helpottaa työntekijöiden toimintaa, koska toiminnan perusteessa dataan, voidaan johdon sekä muun henkilöstön tasolla suorittaa päätöksiä paljon nopeammin. Kyselylomakkeen perusteella huomattiin yhtiön työntekijöiden luottavan dataan ja sen pohjalta tehtyihin päätöksiin. Lisäksi vastaukset kertoivat työntekijöiden mieltävän datan yhtiön tärkeäksi resurssiksi. Data onkin uusi pääoman muoto, jota pidetään talouden tämän hetken tuottavimpana resurssina. Talousyksiköiden työntekijän uskoivat myös organisaatiota soiseen oppimiseen. Vastaajat olivat vahvasti sitä mieltä, että yhtiö kykenee hyödyntämään uutta ja vanhaa tietoa organisaatiossaan hyvin.

Lineaarisen regressiomallin tulosten perusteella teknologiset kyvykkyudet selittivät huonoiten yksilön suorituskykyä. Kyselytutkimukseen vastanneet olivat vahvasti sitä mieltä, ettei yhtiössä jaeta analyttistä tietoa tarpeeksi, eivätkä yhtiön hyödyntämät järjestelmät tue tätä. Vaikka tutkimustulosten perusteella yhtiön talousyksiköissä analysoidaan dataa ja tehdään päätöksiä sen perusteella, ei tieto ilmeisesti leviä jokaisen työntekijän tietoisuuteen. Varsinaiset analyttiset näkemykset ja tiedot saattavat jäädä vain niiden parissa työskentelevien keskuuteen.

Tutkimuksen vastausprosentti jäi melko alhaiseksi, joten tuloksia ei voi välttämättä yleistää yhtiön talousyksiköille. Lisäksi tulokset eivät olleet täysin yksiselitteisiä tutkimuksen luotettavuuteen vaikuttavien tekijöiden vuoksi, ja osa tuloksista olikin ristiriidassa aiempiin tutkimuksiin nähden. Tuloksiin on siis suhtauduttava varauksella ja tarkasteltava kriittisesti.

Jatkotutkimuksia aiheesta voisi suorittaa tiedon jakoon liittyen. Koska tutkittavan yhtiön talousyksiköiden työntekijät olivat sitä mieltä, ettei analyttistä tietoa jaeta organisaatiossa tarpeeksi, olisi syytä tarkastella syitä sen takana. Lisäksi osaavan henkilökunnan tiedon ja osaamisen jakamisesta voisi tarkastella syvemmin ja selvittää siihen liittyvän hiljaisen tiedon merkitystä. Jatkotutkimuksissa voisi laajentaa tämän tutkimuksen tarkastelua myös selittävien muuttujien yhteyksien välille, jotta nähtäisiin miten tekijät vaikuttavat toisiinsa. Lisäksi koko tutkimuksen voisi suorittaa suuremmalle kohderyhmälle, esimerkiksi koko organisaatiolle tai monille eri yhtiöille, jotta vastaajia saataisiin enemmän ja tulokset olisivat yleistettävissä.

## LÄHDELUETTELO

Akter, S., Wamba, S.F., Gunasekaran, A., Dubey, R. & Childe, S.J., (2016) How to improve firm performance using big data analytics capability and business strategy alignment?

Aral, S., Brynjolfsson, E. & Marshall, V.A. (2012) Information, Technology, and Information Worker Productivity. *Information Systems Research*, 23(3), pp. 849-867

Basu, S. K. (2015) The Leader's Role in Managing Change: Five Cases of Technology-Enabled Business Transformation. *Global Business & Organizational Excellence*, 34(3), pp. 28-42.

Busse, S. (2012) Creating a Data-Driven Culture. *Risk Management*, 59(3), pp. 12.

Carlsson, C. (2018) Decision analytics—Key to digitalisation. *Information Sciences*. 460-461

Chae, H., Koh, C.E. & Prybutok, V.R. (2014) Information Technology Capability and Firm Performance: Contradictory Findings and their Possible Causes. *MIS Quarterly*, 38(1), pp. A14

Cooper, C. L. & Nelson, D. (2007) *Positive Organizational Behavior*. London, SAGE Publications Ltd

Corea, F. (2016) *Big Data Analytics: A Management Perspective*. Cham: Springer International Publishing

Dalkir, K. (2005) *Knowledge Management In Theory And Practice*. Oxford, Elsevier

Davenport, T. H. & Harris, J. (2007) *Analysoi ja voita*. Helsinki, Talentum

Davenport, T. H. & Prusak, L. (2000) *Working knowledge: How organizations manage what they know*. Boston (MA): Harvard Business School Press

Farrell, A., Goh, J., Kahle, R., Shackell, M. & White, B. (2017) When Managers Make Emotional Business Decisions. *Strategic Finance* [Verkkodokumentti] [Viitattu 2.11.] Saatavilla <https://sfmagazine.com/post-entry/september-2017-when-managers-make-emotional-business-decisions/>

Frisk, J.E. (2017) Improving the use of analytics and big data by changing the decision-making culture. *Management Decision*, 55(10), pp. 2074-2088.

Grable, J.E. & Lyons, A.C. (2018) An Introduction to Big Data. *Journal of Financial Service Professionals*, 72 (5), pp. 17-20.

Grover, V., Chiang, R.H.L., Liang, T. & Zhang, D. (2018) Creating Strategic Business Value from Big Data Analytics: A Research Framework. *Journal of Management Information Systems*, 35(2), pp. 388-423.

Guadagnoli, E. & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological bulletin*, 103, 265-275

Gupta, B. (2016) *Interview Questions in Business Analytics*. Berkeley, CA: Apress.

Gupta, M. & George, J.F. (2016) Toward the development of a big data analytics capability. *Information & Management*, 53, pp. 1049-1064

Hair, J., Rolph, E., Ronald, L. & Wiliam, C. (1998) *Multivariate Data Analysis*. 5. p. New Jersey, Upper Saddle River

Hill, R. C., Griffiths, W. E. & Lim, G. C. (2012) *Principles of Econometrics*. 4. p. Wiley & Sons

Hirsijäri, S. & Hurme, H. (2001) *Tutkimushaastattelu. Teemahaastattelun teoria ja käytäntö*. Helsinki, Helsinki University Press

Hirsijäri, S., Remes, P. & Sajavaara, P. (2013) Tutki ja kirjoita. 15.-17. Helsinki, Kustannusosakeyhtiö Tammi

Ikäheimo, S., Laitinen E. K., Laitinen, T. & Puttonen, V. (2014) Yrityksen taloushallinto tänään. Vaasa, Vaasan yritysinformaatio Oy

Jiménez-Jiménez, D. & Sanz-Valle, R. (2011) Innovation, organizational learning, and performance. *Journal of Business Research*, 64. pp. 408-417

Laitinen, E. (2003) Yritystoiminnan uudet mittarit. Jyväskylä. Talentum Media Oy

Lam, S.S.K., Chen, X. & Schaubroeck, J. (2002) Participative Decision Making and Employee Performance in Different Cultures: the Moderating Effects of Allocentrism/idiocentrism and Efficacy. *Academy of Management Journal*, 45(5), pp. 905-914.

Lavalle, S., Lesser, E., Shockley, R., Hopkins, M. S. & Kruschwitz, N. (2011) Big Data, Analytics and the Path From Insights to Value. *MIT Sloan Management Review*, 52(2), pp. 21-32

Lind, P. (2015) *Monitoring business performance: models, methods, and tools*. New York, Routledge.

Markkula, T. & Syväniemi, A. (2015) *Analytiikkamatka datasta tietoon ja tiedolla johtamiseen*. Suomen Liikekirjat

McGuire, T., Manyika, J. & Chui, M. (2012) Why Big Data is the New Competitive Advantage. *Ivey Business Journal*. pp. 1.

McNally, J. S. (2018) Business Transformation: No Pain, No Gain? *Strategic Finance* [Verkkodokumentti] [Viitattu 2.11.] Saatavilla <https://sfmagazine.com/post-entry/november-2018-business-transformation-no-pain-no-gain/>



Metsämuuronen, J. (2009) Tutkimuksen tekemisen perusteet ihmistieteissä. Jyväskylä, International Methelp ky

Newell, S., Robertson, M., Scarbrough, H., & Swan, J. (2009) Managing Knowledge Work And Innovation. 2.p. Palgrave Macmillan

O'Donnell F. J. & DUFFY, A.H.B. (2005) Design Performance. London: Springer-Verlag London Limited

Puusa, A., Reijonen, H., Juuti, P. & Laukkanen, T. (2014) Akatemiasta markkinapaikalle. Johtaminen ja markkinointi aikansa kuvina. 4. p. Helsinki, Talentum

Runkler, T. (2016) Data Analytics. Models and Algorithms for Intelligent Data Analysis. 2. p. Springer Vieweg

Reychav, I. & Weisberg, J. (2009) Good for workers, good for companies: How knowledge sharing benefits individual employees. Knowledge & Process Management, 16(4), pp. 186-197.

ROUNDS, M. (2018) Good business metrics drive responsive IT: Are your IT measurement metrics still focused on efficiency and not what is meaningful to the business? If so, it's time to make a change. CIO, , pp. 19-21

Salo, I. (2013) Big Data. Tiedon vallankumous. Jyväskylä, Docendo Oy

Salo, I. (2014) Big Data & Pilvipalvelut. Jyväskylä, Docendo Oy

Shmueli, G., Bruce, P.C. & Patel, N.R. (2016) Data mining for business analytics: concepts, techniques, and applications with XLMiner. Third edition edn. Hoboken: John Wiley & Sons

Smith, D. & Driscoll T. (2017) Growing up in Analytics. Strategic Finance [Verkkodokumentti] [Viitattu 2.11.] Saatavilla <https://sfmagazine.com/post-entry/july-2017-growing-up-in-analytics/>

Sparrow J (2012) The Culture Builders. Leadership Strategies for Employee Performance. Burlington, Gower Publishing Limited

Sujatha, R. & Krishnaveni, R. (2018) Knowledge creating ba as a determinant of work performance of employees: An empirical analysis among pump manufacturing firms in South India. *Asia Pacific Management Review*, 23(1), pp. 45-52

Tabouli, E.M.A., Habtoor, N., A. & Nashief, M. (2016) The impact of human resources management on employee performance: Organizational commitment mediator variable. *Asian Social Science*, 12(9), pp. 176-192.

Ukko, J., Karhu, J., Pekkola, S., Rantanen, H. & Tenhunen, J. (2007) Suorituskyky nousuun! Helsinki, Tykes

Vilkka, H. (2007) Tutki ja mittaa. Määrällisen tutkimuksen perusteet. Helsinki, Kustannusosakeyhtiö Tammi

Virtainlahti, S. (2009) Hiljaisen tietämyksen johtaminen. Helsinki, Talentum

Wong, A. & Sixl-Daniell, K. (2017) Examining the Effectiveness of Corporate E-Learning in Global Talent Management. *International Journal of Advanced Corporate Learning*, 10(2), pp. 4-18.

Xiaojun, Z. (2017) Knowledge Management System use and Job Performance: a Multilevel Contingency Model. *MIS Quarterly*, 41(3), pp. A5.

Zygmunt, C. & Smith, M. R. (2014) Robust factor analysis in the presence of normality violations, missing data, and outliers: Empirical questions and possible solutions. *The Quantitative Methods for Psychology*, 10.

## LIITTEET

### Liite 1. Kyselylomake

<b>Efficiency of the Work</b>	
<b>EW1</b>	I feel dedication, seriousness and ability to take responsibility
<b>EW2</b>	I have professional skill or professionalism and technical knowledge required to carry out the work efficiently
<b>EW3</b>	I do my work according to specific policies and procedures.
<b>EW4</b>	I feel satisfied with the work I do in the company.
<b>Planning the work</b>	
<b>PW1</b>	Planning the work before starting its implementation contributes to setting the goals that need to be achieved
<b>PW2</b>	I have the ability to plan my work and its accomplishment according to the planned schedule.
<b>PW3</b>	Planning the work before starting its implementation gives me a sense of comfort.
<b>PW4</b>	Planning the work before starting its implementation increases my ability to focus on the completion of the work assigned automatically
<b>Creativity and Innovation</b>	
<b>CI1</b>	I am careful or keen to make changes in the working methods of each period.
<b>CI2</b>	I stay away from repeating what others do in solving work-related problems.
<b>CI3</b>	I feel bored of repeating the same procedures in doing the work.
<b>CI4</b>	I have the ability to put forward ideas and solutions rapidly to face work-related problems.

**CI5** I have the ability to express my thoughts fluently and freely.

### **Making effort**

**ME1** Feeling proud of the work represents a motivation for me to make extra efforts.

**ME2 \*** I have the desire and willingness to work outside official working hours for fast delivery.

**ME3 \*** The company is keen on providing additional benefits to employees to motivate them to make more efforts.

### **Analytical Skills**

**TS1** Our organization provides necessary training to our employees related to data and analytics

**TS2** We hire new employees that already have the data analytics skills

**TS3** Our data analytics staff has the right skills to accomplish their jobs successfully

**TS4** Our data analytics staff has suitable education to fulfil their jobs

**TS5** Our data analytics staff is well trained

### **Data-driven Culture**

**DD1** We consider data as an asset

**DD2** We base most of the decisions on data rather than instinct

**DD3 \*** We are willing to override our own intuition when data contradict our viewpoints

**DD4** We continuously assess our strategies and take corrective action in response to the insights obtained from data

**DD5** We are constantly encouraged to make decisions based on data

## **Technology capability**

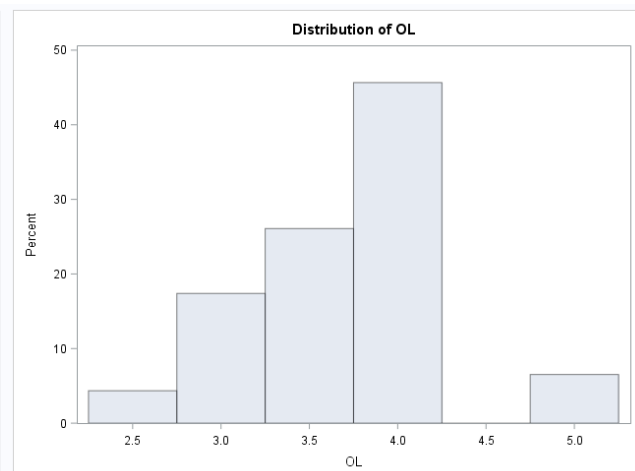
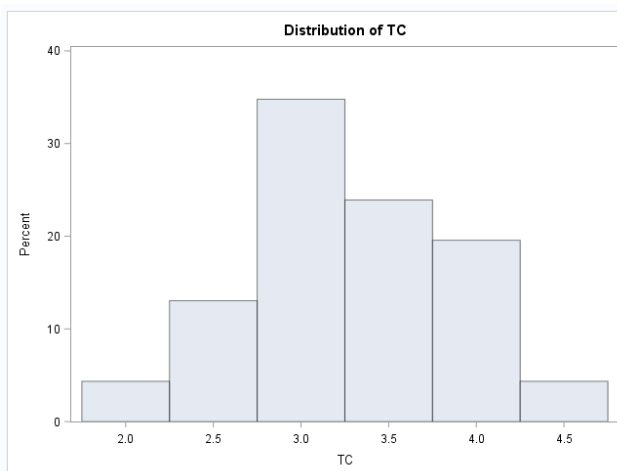
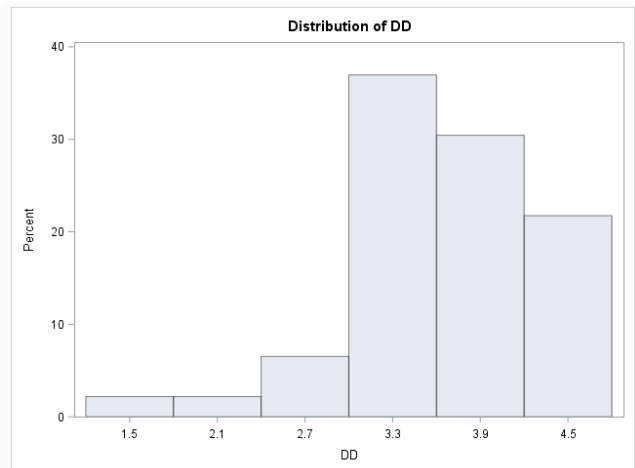
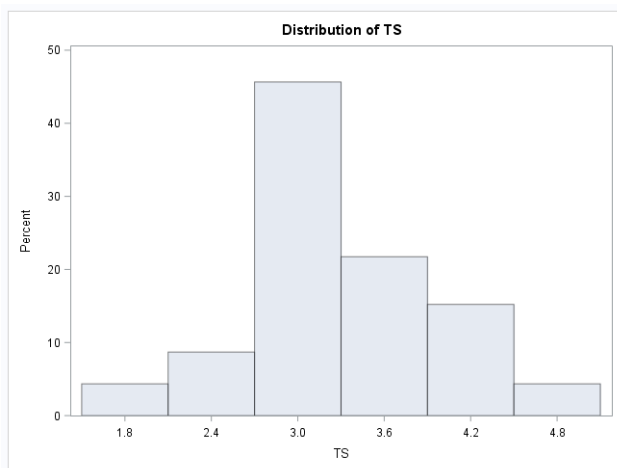
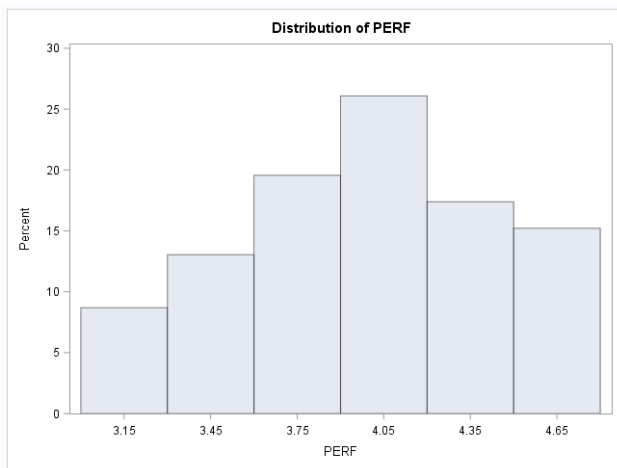
- TC1** All offices (remote, branch and mobile) are connected to the central office for analytical use
- TC2** There are no identifiable communication bottlenecks when sharing analytics insight
- TC3** Software applications can be easily connected to multiple analytics platforms
- TC4** Analytics-driven information is shared seamlessly across our organizations, regardless of the location.
- TC5** We have explored or adopted different data visualization tools
- TC6** We have explored or adopted cloud-based (or internet connected) services for processing data

## **Intensity of Organizational Learning**

- OL1** We are able to search for new and relevant knowledge
- OL2** We are able to reach new and relevant knowledge
- OL3** We are able to adopt relevant knowledge
- OL4** We are able to use relevant knowledge
- OL5** We have made concerted efforts for the exploitation of existing competencies and exploration of new knowledge

*\* väittämät poistuivat tarkastelusta tilastollisten testien perusteella*

## Liite 2. Summamuuttujien jakaumat



Liite 3. Yksittäisten väittämien perustiedot

Variable	Label	Mean	Std Dev	Minimum	Maximum	N
TS1	TS1	3.3043478	0.8658859	1.0000000	5.0000000	46
TS2	TS2	3.2173913	0.8670011	1.0000000	5.0000000	46
TS3 *	TS3	3.3913043	0.7739727	2.0000000	5.0000000	46
TS4	TS4	3.3043478	0.7263169	2.0000000	5.0000000	46
TS5	TS5	3.3043478	0.6623046	2.0000000	5.0000000	46
DD1	DD1	3.9130435	0.7838959	2.0000000	5.0000000	46
DD2	DD2	3.8043478	0.7780193	1.0000000	5.0000000	46
DD3 *	DD3	3.4347826	0.7789501	1.0000000	5.0000000	46
DD4	DD4	3.3260870	0.8957765	1.0000000	5.0000000	46
DD5	DD5	3.4347826	0.7498792	2.0000000	5.0000000	46
TC1	TC1	3.6086957	0.7447076	2.0000000	5.0000000	46
TC2	TC2	2.9782609	0.7742848	1.0000000	5.0000000	46
TC3	TC3	3.0000000	0.8164966	1.0000000	5.0000000	46
TC4	TC4	2.9347826	0.9286112	1.0000000	5.0000000	46
TC5	TC5	3.4130435	0.7767764	2.0000000	5.0000000	46
TC6	TC6	3.5000000	0.8096639	2.0000000	5.0000000	46
OL1	OL1	3.8695652	0.6866568	2.0000000	5.0000000	46
OL2	OL2	3.7608696	0.6389151	2.0000000	5.0000000	46
OL3	OL3	3.6086957	0.7447076	2.0000000	5.0000000	46
OL4	OL4	3.6739130	0.6684758	2.0000000	5.0000000	46
OL5	OL5	3.5434783	0.6898155	2.0000000	5.0000000	46

\* väittämät poistuivat tarkastelusta tilastollisten testien perusteella



# Liite 4. Residuaalikuvaajat

