



Satakunnan ammattikorkeakoulu  
Satakunta University of Applied Sciences

HANNU HATTUNEN

# **AI-tiliöinnin vaikutus tuottavuuteen Palkeiden ostolaskujen käsittelyssä**

LIIKETALouden TUTKINTO-OHJELMA  
2024

## TIIVISTELMÄ

Hattunen, Hannu: AI-tiliöinnin vaikutus tuottavuuteen Palkeiden ostolaskujen käsittelyssä  
Opinnäytetyö, AMK  
Liiketalouden tutkinto-ohjelma  
Helmikuu 2024  
Sivumäärä: 42

Tässä opinnäytetyössä kartoitettiin Snowfox.AI-ostolaskutekoälyn vaikutusta Palkeiden ostolaskujen käsittelyn tuottavuuteen. Työn toimeksiantajana toimi Valtion talous- ja henkilöstöhallinnon palvelukeskus Palkeet, joka tuottaa talous- ja henkilöstöhallinnon palveluja valtion organisaatioille. Tavoitteena oli selvittää, minkälaista osaa ostolaskujen käsittelyn tuottavuudesta keväällä 2022 käyttöönotettu Snowfox.AI selittää.

Opinnäytetyön tutkimusmenetelmänä käytettiin kvantitatiivista tutkimusta. Tutkimus oli luonteeltaan dokumenttianalyysi, eli työssä tutkittiin olemassa olevaa aineistoa. Tutkimusaineistona käytettiin Snowfox.AI:stä sekä Opiferus Talous-suunnittelu-järjestelmästä muodostettua aineistoa. Tilastollisena menetelmänä käytettiin lineaarista regressioanalyysiä, jonka avulla tutkittiin Snowfox.AI:n onnistuneiden tiliöintiennusteiden vaikutusta Palkeiden ostolaskujen käsittelyn tuottavuuteen.

Aineistoa oli käytettävissä Snowfox.AI:n käyttöönotosta opinnäytetyön toteutukseen saakka, yhteensä 19 kuukauden ajalta. Tarkasteltavana oli kaikki 64 Palkeiden asiakasvirastoa, joilla Snowfox.AI on käytössä. Käytössä ollut aineisto oli täydellistä Snowfox.AI:n sekä Palkeiden tuottavuuslaskelmien osalta, joten tutkimus voitiin tehdä kokonaistutkimuksena.

Regressioanalyysin tuloksena saatiin Snowfox.AI:n tuottavuutta selittävät lukemat, joita voidaan hyödyntää esimerkiksi tulevien tuottavuuslaskelmien vertailukohtana. Tulevien tulosten avulla voidaan todeta, minkälaisia vaikutuksia ostolaskutekoälyn kehittämisellä on saatu aikaiseksi.

Avainsanat: tekoäly, tuottavuus, taloushallinto, lineaarinen regressioanalyysi, ostolasku, Palkeet

## ABSTRACT

Hattunen, Hannu: The effect of Snowfox.AI on productivity in the processing of the Palkeet's purchase invoices

Bachelor's thesis

Business Administration

February 2024

Number of pages: 42

The thesis examined the effect of Snowfox.AI on productivity in the processing of the Palkeet's purchase invoices. The client of the work was The Finnish Government Shared Services Centre for Finance and HR (Palkeet). Palkeet provides financial and HR administration services to the government. The goal was to find out what part of the productivity of processing purchase invoices Snowfox.AI explains.

Quantitative research was used as the research method of the thesis. The nature of the research was document analysis, so the work examined existing material. The data generated from Snowfox.AI and the Opiferus Taloussuunnittelu system were used as research data. Linear regression analysis was used as a statistical method, which was used to study the effect of Snowfox.AI's successful invoicing prediction on the productivity of processing purchase invoices.

The data was available from the start date of Snowfox.AI at Palkeet until the implementation of my thesis, for a total of 19 months. The subject of the research was 64 Palkeet's customer agencies that uses Snowfox.AI. The data used was perfect, so the study could be done as a census study.

As a result of regression analysis are the readings that explain the productivity of Snowfox.AI, which can be used as a comparison of future productivity calculations. With the help of the future results, it can be stated what kind of effects the development of artificial intelligence invoicing has achieved.

Keywords: artificial intelligence, productivity, accounts payable, financial management, linear regression analysis, Palkeet

# SISÄLLYS

1 JOHDANTO .....	5
2 TYÖN TAVOITE JA KÄYTETTÄVÄT MENTELMÄT .....	6
2.1 Toimeksiantajan esittely .....	6
2.2 Opinnäytetyön tavoite ja tutkimuskysymys .....	7
2.3 Tutkimusmenetelmä .....	9
2.4 Dokumenttianalyysi .....	11
2.5 Lineaarinen regressioanalyysi .....	12
3 ÄLYKÄS TALOUSHALLINTO JA TUOTTAVUUS .....	14
3.1 Kohti älykästä taloushallintoa .....	14
3.2 AI älykkäässä taloushallinnossa .....	16
3.3 Tuottavuus julkishallinnossa .....	21
4 REGRESSIOMALLIN MUODOSTAMINEN JA TULOKSET .....	23
4.1 Lineaarisen regressioanalyysin perusteet .....	23
4.2 Tutkimusaineiston koostaminen .....	26
4.3 Tutkimusaineiston yhdistely .....	27
4.4 Poikkeavat havainnot ja aineiston muokkaus .....	28
4.5 Regressioanalyysin tulokset .....	31
4.6 Validiteetti ja reliabiliteetti .....	35
5 JOHTOPÄÄTÖKSET .....	36
6 POHDINTA .....	38
LÄHTEET .....	40

## 1 JOHDANTO

Älykäs taloushallinto on 2020-luvun taloushallinnon kasvava trendi. Koneoppivat tekoälyratkaisut yleistyvät ja integroituvat taloushallinnon ohjelmistoihin. Tekoälyratkaisut soveltuvat hyvin taloushallinnon ohjelmistoihin, koska käsiteltävät aineistot ovat pääosin digitaalisia, perustuvat loogisiin sääntöihin ja ovat matemaattisesti mallinnettavissa. Tekoälyratkaisut vähentävät aineistojen rutiininomaista manuaalista käsittelyä, vapauttavat ihmisen tekemää työtä asiantuntijatehtäviin sekä parantavat organisaation tuottavuutta.

Opinnäytetyöni toimeksiantajana toimi Valtion talous- ja henkilöstöhallinnon palvelukeskus Palkeet, joka tuottaa talous- ja henkilöstöhallinnon palveluita valtion organisaatioille. Palkeet hyödyntää toiminnoissaan automatiikkaa ja robotiikkaa ja yksi tulevien vuosien kehittämisen painopisteistä Palkeilla on taloushallinnon automatisoinnin kehittäminen.

Tämän opinnäytetyön tarkoituksena on kartoittaa Snowfox.AI-ostolaskutekoälyn vaikutusta Palkeiden ostolaskujen käsittelyn tuottavuuteen. Palkeilla käsitellään vuosittain noin 1,3 miljoonaa ostolaskua, joten pienelläkin tuottavuuden parantamisella voidaan saada näkyviä muutoksia. Snowfox.AI luo Palkeiden ostolaskujen historiadatan perusteella tiliöintiennusteen ostolaskulle, jolloin laskun käsittelijän tehtäväksi jää ennusteen tarkistaminen ja mahdolliset korjaustoimenpiteet. Snowfox.AI:n ennusteiden onnistumisprosenttia ja sen käsittelemien laskujen määrää on seurattu Palkeilla käyttöönotosta saakka, mutta tekoälyn vaikutusta tuottavuuteen ei ole aiemmin mitattu.

Kustannustehokkuus on yksi Palkeiden perustehtävistä, joten tuottavuuksien ja osatuottavuuksien mittaaminen on tärkeä osa sen toteutumista. Työn tulokset antavat Palkeille uuden näkökulman tuottavuuden tarkasteluun ja tuloksia voidaan hyödyntää mm. vertailukohtana tulevissa tuottavuuden tarkasteluissa.

## 2 TYÖN TAVOITE JA KÄYTETTÄVÄT MENTELMÄT

### 2.1 Toimeksiantajan esittely

Valtion taloushallinnolla on merkittävä rooli valtion varojen asianmukaisen käytön varmistamisessa sekä hyvän hallinnon menettelyjen ylläpitämisessä. Keskeisimpiä toimijoita valtion taloushallinnon järjestämisessä ovat Valtiovarainministeriö, Valtiokonttori, eri hallinnonalojen ministeriöt ja virastot sekä Valtion talous- ja henkilöstöhallinnon palvelukeskus Palkeet. (Valtiovarainministeriö, 2023b).

Valtiovarainministeriö vastaa taloushallinnon säädöksistä, konsernilinjauksista sekä strategisesta ohjauksesta. Valtiokonttori ohjaa eri toimijoita järjestämään kirjanpidon ja muun laskentatoimen yhtenäisellä tavalla. Virastojen taloushallinto vastaa virastojen talouden suunnittelusta sekä seurannasta. Valtion talous- ja henkilöstöhallinnon palvelukeskus Palkeet tuottaa valtion virastojen talous- ja henkilöstöhallinnon tuki- ja asiantuntijapalveluita. (Valtiovarainministeriö, 2023b).

Palkeiden toimintaa säätelee laki Valtion talous- ja henkilöstöhallinnon palvelukeskuksesta 179/2019. Palkeiden tehtävänä on:

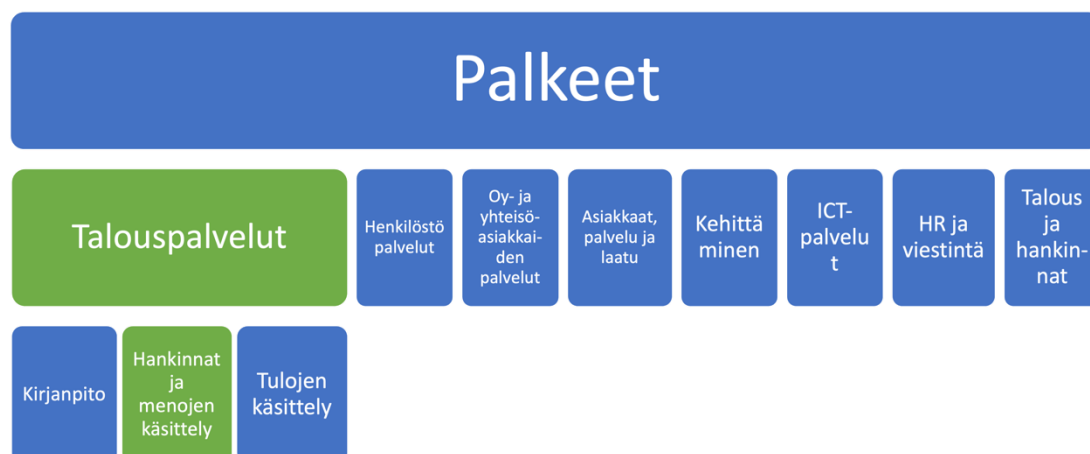
1. tuottaa keskitettyjä taloushallintotehtäviä siten kuin Valtiokonttori on valtion talousarviosta annetun lain nojalla palvelukeskuksen tehtäväksi määrännyt;
2. tuottaa talous- ja henkilöstöhallinnon palveluja sekä muita vastaavia hallinnon tuki- ja asiantuntijapalveluja siten kuin niistä on palvelusopimuksissa sovittu;
3. tuottaa talous- ja henkilöstöhallinnon analysointi- ja raportointipalveluja palvelukeskuksen asiakkaille valmistelun ja päätöksenteon tueksi. (Laki Valtion talous- ja henkilöstöhallinnon palvelukeskuksesta, 179/2019, 1 §).

Tällä hetkellä Palkeilla työskentelee noin 650 henkilöä Joensuun, Helsingin, Hämeenlinnan, Mikkelin sekä Porin toimipaikoissa (Palkeet, 2023b).

Taulukko 1. Palkeiden vuoden 2022 tunnuslukuja (Palkeet, 2023b).

Liikevaihto	59,5 miljoonaa euroa
Ostolaskut	1,2 miljoonaa kpl
Myyntilaskut	0,6 miljoonaa kpl
Palkkalaskelmat	1,2 miljoonaa kpl
Matkalaskut	0,3 miljoonaa kpl
Tilinpäätös	70 kirjanpitoyksikölle ja rahastolle ja 8 osakeyhtiö- ja yhteisöasiakkaalle

Palkeiden organisaatio jakaantuu kahdeksaan yksikköön (Kuvio 1). Palkeiden talouspalveluihin kuuluu kirjanpito, hankinnat ja menojen käsittely sekä tulojen käsittely. Opinnäytetyöni aihe sijoittuu talouspalveluihin, hankintojen ja menojen käsittelyyn alle.



Kuvio 1. Palkeiden organisaatiokaavio.

## 2.2 Opinnäytetyön tavoite ja tutkimuskysymys

Opinnäytetyöni tavoitteena on kartoittaa tilastollisella menetelmällä keväällä 2022 käyttöön otetun Snowfox.AI-ostolaskutekoälyn vaikutusta Palkeiden ostolaskujen käsittelyn tuottavuuteen. Tavoitteena on tutkia, kuinka suurta osaa Palkeiden ostolaskujen käsittelyn tuottavuudesta tekoäly selittää. Tekoälyn onnistumisprosenttia ja ostolaskujen käsittelyn automaatioastetta mitataan Palkeilla, mutta nämä luvut eivät kerro vaikutusta kokonaistuottavuuteen. Opinnäytetyössäni mitataan osatuottavuus Snowfox.AI:n osalta sen käyttöönotosta tutkimusajankohtaan saakka. Snowfox.AI:n toiminta kehittyy jatkuvasti

sen saadessa lisää oppimisdataa, joten opinnäytetyön tuloksia voidaan hyödyntää esimerkiksi tulevien tuottavuusmittausten vertailukohtana, jolloin nähdään minkälaisia vaikutuksia Snowfox.AI:n ennusteiden kehityksellä on ollut tuottavuuden kehitykseen.

Palkeilla käsitellään vuosittain noin 1,2 miljoonaa ostolaskua. Määrät ovat suuria, joten pienilläkin tuottavuutta parantavilla tekijöillä voi olla merkittävä vaikutus tuottavuuteen. Kustannustehokkuus on osa Palkeiden perustehtävää: ”Palkeiden tavoitteena on edistää julkishallinnon toimivuutta tarjoamalla laadukkaita ja kustannustehokkaita talous- ja henkilöstöhallinnon tuki- ja asiantuntijapalveluja” (Palkeet 2023a).

Valtionhallinnossa työn tuottavuuden nostoa on tavoiteltu vuoden 2002 talousarvioneuvotteluissa valmistellun tuottavuusohjelman avulla. Palkeet aloitti toimintansa vuonna 2010 vastatakseen tarpeeseen tehostaa valtion talous- ja henkilöstöhallintoa. Silloin arvioitiin yhteisen talous- ja henkilöstöhallinnon tietojärjestelmän, palvelukeskustoimintamallin ja yhteisten seurantakohteiden ja prosessien nostavan valtion talous- ja henkilöstöhallinnon tuottavuutta 30–40 prosenttia aikaisempaan hajautettuun käsittelyyn verrattuna (Tarkka, 2010, s. 3; Palkeet, 2023a). Vuonna 2022 Palkeet saavutti vuosittaisen kokonaistuottavuuden tavoitteen. Ostolaskujen käsittelyssä parannusta edellisvuoteen oli 15,8 prosenttia (Valtiovarainministeriö, 2023a, s. 2).

Tutkin tuottavuutta älykkään taloushallinnon viitekehyksessä, johon myös Palkeilla vuonna 2022 käyttöön otettu Snowfox.AI-ostolaskutekoäly linkittyy. Aihe on Palkeiden osalta ajankohtainen, koska yksi Palkeiden vuosien 2022–2026 kehittämisen painopisteistä on taloushallinnon automaation kehittäminen (Palkeet, 2023c). Tämän opinnäytetyön tavoitteena on tuottaa lisätietoa ostolaskutekoällyn vaikutuksesta tuottavuuteen esimerkiksi tulevien vuosien tuottavuustavoitteiden saavuttamista varten.

Työni tutkimuskysymys on:

Miten Snowfox.AI-ostolaskutekoäly vaikuttaa Palkeiden ostolaskujen käsittelyn tuottavuuteen?



## 2.3 Tutkimusmenetelmä

Tieteellinen tutkimus voi olla teoreettista, ns. kirjoituspöytä tutkimusta tai havainnoivaa, eli empiiristä tutkimusta. Empiirinen tutkimus, jota opinnäytetyöni edustaa, perustuu teoreettisen tutkimuksen perusteella kehitettyihin menetelmiin. Empiirisessä tutkimuksessa testataan, toteutuuko jokin teoriasta johdettu oletamus käytännössä. Tutkimusongelmana voi olla myös jonkin ilmiön syiden selvittäminen. (Heikkilä, 2014, s.12). Empiirisessä tutkimuksessa pyritään löytämään yksittäistapausten perusteella yleisiä lainalaisuuksia sekä säännönmukaisuuksia (Valli, 2015, s.10).

Empiirinen tutkimus voidaan jakaa kvantitatiivisiin, eli tilastollisiin tai määrällisiin tutkimuksiin sekä kvalitatiivisiin, eli laadullisiin tutkimuksiin. Kvantitatiivisen tutkimuksen avulla selvitetään lukumääriin ja prosenttiosuuksiin liittyviä kysymyksiä. Tutkimustuloksia voidaan havainnollistaa taulukoiden ja kuvioiden avulla. Tutkimustapa edellyttää riittävän suurta ja edustavaa otosta ollakseen luotettava. (Heikkilä, 2014, s.15).

Tilastollista mallia muodostaessa tarkastellaan reaali maailmasta kerättyä havaintoaineistoa, joka koostuu useista havaintoyksiköistä. Pyrkimyksenä on löytää mahdolliset johdonmukaisuudet ja rakenteet tutkittavien ominaisuuksien väliltä. Mikäli sama ilmiö toistuu, testataan ilmiön tilastollinen yleistettävyyden ja voidaanko asia yleistää koskemaan koko perusjoukkoa (Valli, 2015, s.11; Heikkilä, 2014, s.32).

Perusjoukon ominaisuuksia tutkittaessa on ratkaistava, tehdäänkö tutkimus kokonaistutkimuksena, jossa populaation jokainen otantayksikkö otetaan tarkasteltavaksi. Tämä on harvoin mahdollista, joten usein ainoa vaihtoehtona on suorittaa otantatutkimus, jossa tietty osajoukko voisi edustaa koko populaatiota. (Nummenmaa et al., 2019, s.25–26). Kokonaistutkimus kannattaa tehdä, jos yksiköiden lukumäärä on alle sata, mutta kokonaistutkimus voidaan tehdä suuremmastakin perusjoukosta esimerkiksi mitattavan ominaisuuden suuren vaihtelun vuoksi. Otantatutkimukseen päädytään muun muassa silloin, mikäli

perusjoukko on hyvin suuri, tutkiminen maksaisi liian paljon tai tiedot halutaan mahdollisimman nopeasti. (Heikkilä, 2014, s.31).

Opinnäytetyöni tehtiin kokonaistutkimuksena, koska saatavilla oleva data oli aukotonta ja valmiiksi dokumentoitu Snowfox.AI:n käyttöön otosta saakka. Tarkasteluun valittiin Palkeiden asiakkaat, joilla Snowfox.AI on käytössä. Palkeiden asiakkaista ulkopuolelle jäivät osakeyhtiö- sekä yhteisöasiakkaat, osa rahastoista sekä Puolustusvoimat, jotka toimivat eriyttynä muista asiakkaista.

Kartoittavaa (eksplanatiivista) tutkimusta käytetään, kun aiheesta on vain vähän tai ei lainkaan aiempaa tietoa. Sitä käytetään, kun olemassa oleva tieto ei riitä konkreettisten hypoteesien muodostamiseen. Kartoittava tutkimus voi toimia alustavana tutkimuksena, jonka tavoitteena on hankkia tietoa, herättää ajatuksia ja havaita mahdollisia malleja tutkittavassa aiheessa. Se voi auttaa määrittelemään tutkimuskysymyksiä, kehittämään hypoteeseja ja suunnittelemaan tarkemmin kohdennettuja tutkimuksia tulevaisuudessa. (Abbadia, 2023).

Kartoittava tutkimus voi olla ns. primääritutkimusta (ensisijaista tutkimusta) tai sekundaarista (toissijaista) tutkimusta. Primääritutkimuksessa tutkittava aineisto kerätään suoraan lähteestä esimerkiksi kyselyillä, haastatteluilla, havainnoinnilla tai kokeilla. Sekundääritutkimuksen aineisto on peräisin olemassa olevissa lähteissä, kuten raporteissa, tutkimuksissa tai tietokannoissa. (Abbadia, 2023). Tämän opinnäytetyön aineisto on sekundääriaineistoa, eli tutkimusaineistoa ei ole alun perin kerätty tätä opinnäytetyötä varten, vaan se on Palkeiden järjestelmien muodostamaa dataa. Opinnäytetyön aineisto on muodostettu Snowfox.AI sekä Opiferus Taloussuunnittelu-ohjelmistojen keräämästä datasta.

Kartoittava tutkimustapa on perusteltu työni kannalta, koska Palkeilla järjestelmiin tallentuu paljon tuottavuuden analysointiin sopivaa aineistoa, jota ei suoraan hyödynnetä siihen. Palkeiden Snowfox.AI:n ennusteiden onnistumisprosenttia ja sen käsittelemien laskujen määrää on seurattu käyttöön otosta saakka, mutta tekoälyn vaikutusta tuottavuuteen ei ole aiemmin mitattu.

Työtäni voisi siis kuvailla tilastollisella menetelmällä toteutetuksi kartoittavaksi dokumenttianalyysiksi.

## 2.4 Dokumenttianalyysi

Dokumenttianalyysillä tarkoitetaan olemassa olevan aineiston tutkimista. Dokumenttianalyysiä käytetään tilanteissa, joissa aineistoa ei saada koottua välittömien havaintojen avulla. Aineistoa tutkitaan ja tulkitaan loogisen päättelyn avulla. Dokumenttianalyysin vahvuus on se, että tutkittava materiaali on omassa luonnollisessa ympäristössään, joten menetelmää voidaan hyödyntää myös sellaisen materiaalin analysointiin, jota ei alun perin ole tarkoitettu tutkittavaksi. Dokumenttianalyysin tavoitteena on tietopääoman lisääminen. (Sivonen, 2017, s. 15).

Opinnäytetyöni lähdeaineistoina toimi Palkeiden järjestelmistä muodostettu aineisto. Snowfox.AI:n raportit sisälsivät kuukausittaiset asiakaskohtaiset tiliöityjen laskujen suoritemäärät sekä tiliointien onnistumisprosentit. Tuottavuuteen liittyvät raportit saatiin Opiferus Taloussunnittelu-järjestelmästä, johon tallentuu eri toiminnoille kirjatut työajat ja ostolaskujen suoritemäärät. Tuottavuusraportille muodostettiin eri ostolaskutyypin kuukausittaiset virastokohtaiset suoritemäärät, teholliset henkilötyövuodet sekä tuottavuudet.

Työn toimeksiantajan kanssa sovittiin, että kaikki työhön liittyvä aineisto käsitellään Palkeiden tietokoneella ja tallennetaan joko Palkeiden työasemalle tai Palkeiden sisäiseen verkkoon. Ulkopuolisten pääsy aineistoon käsiksi estyy sillä, että materiaali käsitellään työasemilla, jotka vaativat henkilökortin tai vahvan tunnistautumisen auetakseen. Henkilötietoja sisältävän materiaalin käsittely ei ole tutkimusongelman ratkaisun kannalta tarpeellista eikä sellaista tietoa kerätty opinnäytetyötä varten. Aineiston omistaa Palkeet ja Palkeilla on käyttöoikeus kerättyyn aineistoon. Palkeiden kanssa sovitaan tavoista, joilla tutkimuksen tuloksia voidaan esittää opinnäytetyön kirjallisessa osiossa sekä opinnäytetyöhön liittyvissä seminaareissa. Tulokset tulee esittää siten, ettei asiakas ole tunnistettavissa.

## 2.5 Lineaarinen regressioanalyysi

Kvantitatiiviset tutkimukset voidaan jakaa havainnoiviin ja kokeellisiin tutkimuksiin. Havainnoiva tutkimus voi olla prospektiivista, eli ilmiön kehittymisen seuraamista tulevaisuuteen tai retrospektiivistä, eli ilmiön tutkimista takautuvasti. Näissä tutkija ei manipuloi mitään tutkimusasetelman muuttujaa, vaan tutkii muuttujia niiden omassa ominaisuudessaan. Kokeellisessa tutkimuksessa riippumattomia muuttujia pyritään manipuloimaan, jolloin voidaan tehdä päätelmiä riippumattoman ja riippuvan muuttujan välisistä syy-seuraussuhteista. Poikkileikkaustutkimuksessa tutkitaan ilmiön eri puolia tietyssä ajankohdassa. Kuvaileva tutkimus esittää asiointilan tietyssä hetkenä tai pitkän ajan kuluessa. Case-tutkimuksissa tutkitaan tietty yksittäistapaus mahdollisimman tarkasti ja tehdään sitä koskevat johtopäätökset, toimenpidesuositukset tai yleistyksiset. (Nummenmaa et al., 2019, s.16–17).

Tutkimusta tehdessä on aluksi määriteltävä, mikä joukko on tutkimuksen kohteena. Joukkoa kutsutaan populaatioksi tai perusjoukoksi, joka muodostuu tutkimusyksiköistä. Tilastollisessa tutkimuksessa näitä kutsutaan tilastoyksiköiksi. Kun tilastoyksiköistä kerätään tietoja mittaamalla, kohdistuu mittaaminen muuttuun. Muuttujat ovat siis tutkimuksessa mittauskohteita. Kvantitatiiviset muuttujat voivat olla diskreettejä tai jatkuvia. Diskreetit muuttujat voivat saada vain yksittäisiä arvoja ja jatkuva muuttuja voi saada tietyllä välillä kaikki mahdolliset arvot. (Nummenmaa et al., 2019, s.17–18).

Erilaisia muuttujia tulee mitata eri keinoilla ja eri mitta-asteikoilla. Tilastollisen tutkimuksen mitta-asteikot voidaan jakaa laatuero- eli luokitteluasteikkoon, järjestysasteikkoon, välimatka-asteikkoon sekä suhdeasteikkoon. Laatuero- eli luokitteluasteikolla tilastoyksiköt luokitellaan määrättyihin luokkiin, joiden järjestyksellä ei ole merkitystä. Kukin tilastoyksikkö voi kuulua vain yhteen luokkaan, jolloin kahdesta tilastoyksiköstä voidaan päätellä kuuluvatko ne samaan luokkaan vai ei. Luokitteluasteikon tasoiset mittauksia ovat esimerkiksi sukupuolen, kotikunnan tai veriryhmän mittaukset. Järjestysasteikolla tilastoyksiköt luokitellaan ennalta määrättyihin luokkiin, joilla on yksikäsitteinen järjestys.

Kilpailujen tulokset, joissa on voittaja, toiseksi tullut jne. on yksi esimerkki järjestysasteikosta. (Nummenmaa et al., 2019, s.18).

Välimatka-asteikon mittaustulokset ovat reaalitylukuja ja mitattujen havaintoarvojen etäisyydet on mahdollista määrittellä. Näin ollen mittaustuloksille voidaan myös suorittaa laskutoimituksia. Välimatka-asteikollisilla muuttujilla ei kuitenkaan ole absoluuttista nolapistettä, vaan arvot voivat olla sekä negatiivisia että positiivisia. Suhdeasteikolliset mittaustulokset ovat myös reaalitylukuja, mutta asteikolla on myös absoluuttinen nolapistete. Tällöin mittaustulosten välisten suhteiden laskeminen on mielekäästä. Tilastollisen tutkimuksen mitta-asteikko on valittava huolellisesti, koska sitä ei voi jälkikäteen muuttaa, jos esimerkiksi huomataan tilastollisen analyysimenetelmän vaativan sitä. (Nummenmaa et al., 2019, s.18–19).

Opinnäytetyöni tilastollisena menetelmänä käytän lineaarista regressioanalyysiä. Lineaarilla regressioanalyysillä voidaan mallintaa yhden tai useamman selittävän muuttujan ja yhden selitettävän muuttujan välistä riippuvuutta (Valli, 2015, s. 77). Selittävällä muuttujalla tarkoitetaan muuttujaa, jonka vaikutusta tutkittavaan ominaisuuteen pyritään selvittämään. Selitettävä muuttuja on tutkimuksen pääkohde, johon vaikuttavia tekijöitä, eli selittäjiä etsitään. (Heikkilä, 2014, s. 280). Regressioanalyysin otoskoon on oltava vähintään 50 ja muuttujien tulee olla välimatka- tai suhdeasteikkoisia. Muuttujien tulee olla normaali-jakautuneita ja niissä ei saa olla yksittäisiä poikkeavia arvoja. Selittävät muuttujat eivät saa korreloida voimakkaasti keskenään. (Valli, 2015, s. 77).

Regressiomallille on useita mahdollisia käyttötarkoituksia. Lineaarisen regressiomallin avulla voidaan esimerkiksi autokaupassa selittää tai ennustaa ajettujen kilometrien ja vuosimallin vaikutusta käytetyn auton hintaan. Tällöin ajettujen kilometrit ja vuosimalli olisivat selittäviä muuttujia ja myyntihinta selitettävä muuttuja. Sitä voidaan hyödyntää myös esimerkiksi ennustavana mallina kiinteistönvälityksessä ennustamaan myyntiin tulevien asuntojen hintoja. Regressiomallia voidaan käyttää myös selittävänä mallina, kuten opinnäytetyössäni, jossa pyrin selittämään minkälaisista osaa tuottavuudesta Snowfox.AI-ostolaskutekoäly selittää. (Taanila, 2020, s.1).

Opinnäytetyössäni selitettävä muuttuja on Palkeiden ostolaskujen käsittelyn tuottavuus (kpl / tehollinen HTV). Selittävänä muuttujana toimii sataprosenttisesti oikein AI-tiliöityjen ostolaskujen määrät. Ennen regressioanalyysiä tulee selvittää, voiko muuttujilla olla teoreettisesti mitään yhteyttä keskenään, eli läpäiseekö aineisto järkevyytarkastelun (Nummenmaa et al., 2019, s.237). Opinnäytetyöni muuttujien valinnat perustuvat oletuksiin, että AI-tiliöinnin vaikutus tuottavuuteen on positiivinen ja että korkeampi AI-tiliöityjen laskujen osuus selittää tuottavuutta, eikä esimerkiksi toisinpäin. Järkevyytarkastelun perusteena on oletus, että AI-tiliöinti nopeuttaa ostolaskujen käsittelyä vähentäen ihmisten tekemää manuaalista työtä, joten voidaan olettaa, että sillä on positiivinen vaikutus tuottavuuteen. Valituilla muuttujilla pyritään selvittämään, minkälaista osuutta tuottavuudesta AI-tiliöinnin toiminta selittää.

## 3 ÄLYKÄS TALOUSHALLINTO JA TUOTTAVUUS

### 3.1 Kohti älykästä taloushallintoa

Taloushallintoala on monien muiden alojen joukossa muuttanut muotoaan digitalisaation myötä. Yksi ilmeisimmistä digitalisaation mahdollistamista asioista taloushallinnon alalla oli siirtyminen pois paperitositteiden käsittelystä. Jokaisella lienee vähintään mielikuva stereotyyppisestä toimistotyöstä, jossa tositteita mapitetaan, skannataan ja tulostetaan.

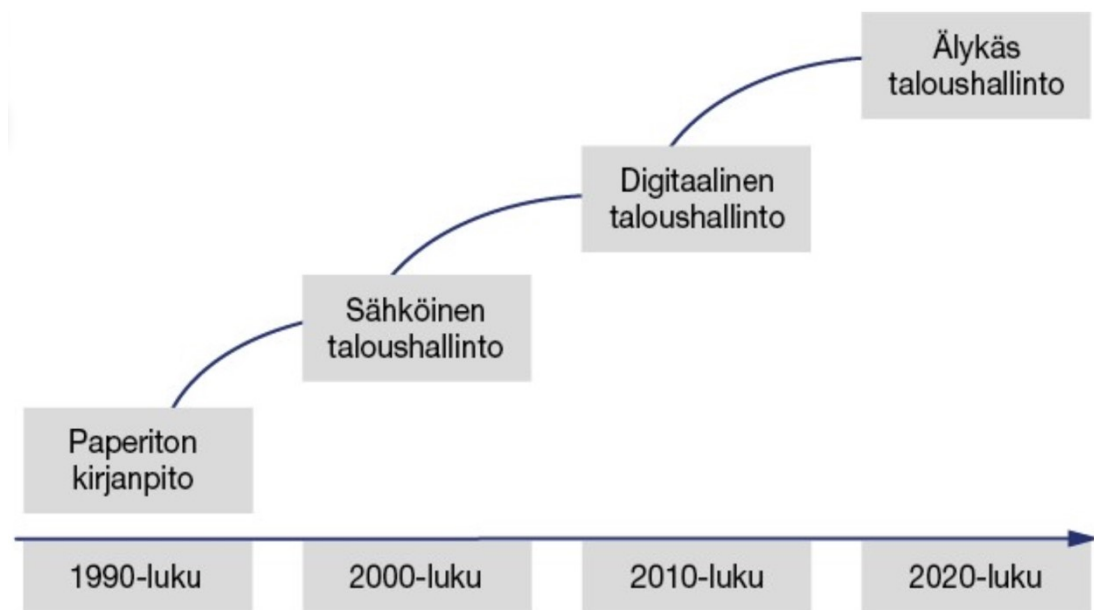
Suomi toimi globaalina edelläkävijänä olemalla ensimmäinen maa, joka siirtyi paperisesta taloushallinnosta sähköiseen. Siirtyminen mahdollistui vuoden 1997 lainsäädännön muutosten myötä. Siihen saakka lain vaatimusten vuoksi taloushallinnon tositteet olivat vähintään arkistoitu paperimuodossa. (Kaarlejärvi & Salminen, 2018, s.11). Paperittoman taloushallinnon uutisoitiin lainsäädännön muutoksen jälkeen tuovan miljardien markkojen vuotuiset säästöt. Lainmuutosten ajan uutisointi oli toiveikasta ja maalasi kuvaa ajasta, jolloin

myyntireskontraa ei enää tarvittaisi ja tilinpäätös voitaisiin tehdä valmiiksi heti tilikauden viimeisenä päivänä, koska kaikki tarvittava tieto on jo koneella. (Ilta-sanomat, 1998). Taloudelliset säästöt ovat eittämättä olleet merkittävät varsinkin byrokraattisesti raskaassa julkishallinnossa. Kaikki lainmuutosten tuomat toiveet eivät tosin ole toteutuneet vielä tähän päivään mennessä.

Paperittomalla kirjanpidolla tarkoitetaan kirjanpidon lakisääteisten tositteiden esitystapaa sähköisessä muodossa (Kaarlejärvi & Salminen, 2018, s.15). Käytännössä paperiton kirjanpito on kuitenkin voinut tarkoittaa tehotonta ja manuaalista tositteiden käsittelyä, kuten papereiden skannaamista sähköiseen muotoon. Sähköisestä taloushallinnosta alettiin puhua vasta myöhemmin, jolloin taloushallintoa tehostettiin tietotekniikkaa, sovelluksia, internetiä sekä erilaisia sähköisiä palveluita hyödyntämällä (Lahti & Salminen, 2014, s.26). Digitaalisesta, tai automaattisesta taloushallinnosta alettiin puhua vasta 2010-luvulla, jolloin taloushallinnon kaikkien tietovirtojen ja käsittelyvaiheiden automatisointi sekä käsittely oli mahdollista digitaalisessa muodossa. Digitaalinen taloushallinto mahdollisti sääntöpohjaisen automaation hyödyntämisen, jolla osa prosesseista voidaan automatisoida erilaisten sääntöjen ja validointien perusteella. (Kaarlejärvi & Salminen, 2018, s.14–16).

Palkeilla taloushallinnon digitalisoituminen näkyy esimerkiksi sääntöpohjaisen automaation ja robotiikan hyödyntämisenä. Robotit tekevät paljon rutiinimaista työtä, kuten aineistojen täsmäyttämistä, laskutusaineistojen sisään lukua ja ostolaskujen perustietojen tarkistamista, joka vapauttaa talouspalveluiden asiantuntijoiden työtä muihin tehtäviin, kuten asiakaspalveluun sekä neuvontaan. Muun muassa ostosta maksuun -prosessi (Palkeilla tarpeesta maksuun -prosessi) on automatiikan myötä mahdollista hoitaa merkittävästi nopeammin kuin ennen ja näin välttää maksujen viivästymiset ja niistä aiheutuvat kustannukset (Kaarlejärvi & Salminen, 2018, s.22). Robotiikkaa hyödynnetään Palkeilla myös rajapintaratkaisuna eri järjestelmien välillä siirtämään tietoa saumattomasti järjestelmästä toiseen. Ihmisen työksi useissa prosesseissa jää vain automatiikan/robotin tekemän työn hyväksyminen ja mahdollisten virhetilanteiden selvittäminen.

2020-luvun trendinä on älykäs taloushallinto, jossa järjestelmät luovat itselleen käsittelysääntöjä, käsittelevät poikkeamatilanteita, analysoivat lopputuotoksia ja ennustavat tulevaa. Järjestelmät korvaavat yhä enemmän ihmistä rutiinimaisissa tehtävissä ja tukevat ihmisten tekemää työtä ongelmanratkaisua vaativissa tehtävissä. Tekoälyä ja koneoppimista hyödynnetään yhä enemmän taloushallinnon ohjelmistoissa, joihin tekoälyn teknologiat alkavat myös integroitua. (Kaarlejärvi & Salminen, 2018, s.17–20). Kuvio 2 havainnollistaa taloushallinnon digitalisaation kehittymistä viime vuosikymmeninä.



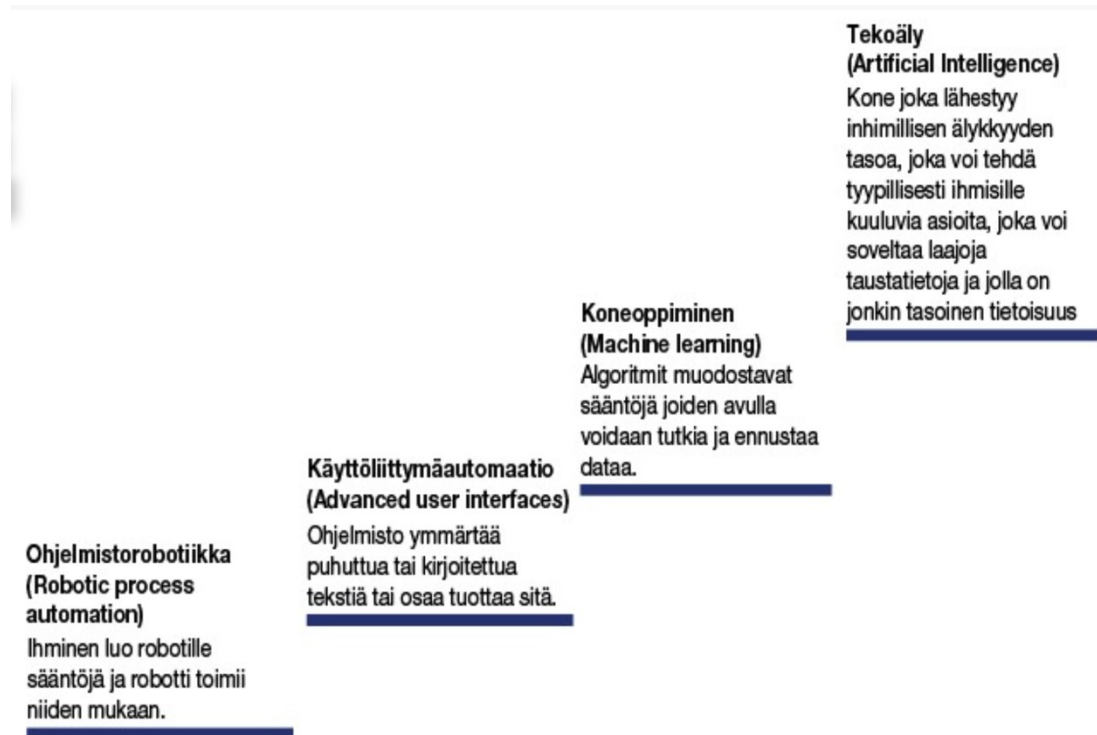
Kuvio 2. Taloushallinnon digitalisoituminen (Kaarlejärvi & Salminen, 2018, s.16)

### 3.2 AI älykkäässä taloushallinnossa

Taloushallinnossa robotiikan ja tekoälyn sovelluksia löytyy yksinkertaisista ohjelmistorobotiikan sovelluksista inhimillistä älykkyyttä lähestyviin teknologioihin. Kuvio 3 selkeyttää yleisesti käytössä olevia termejä. Taloushallinnon prosesseissa hyödynnetyin robotiikan muoto on ohjelmistorobotiikka, joka sopii säännönmukaisiin tehtäviin, joissa käsitellään määrämuotoista tietoa. Kehittyneellä käyttöliittymäautomaatiolla pystytään esimerkiksi muokkaamaan vapaamuotoinen sähköpostiviesti määrämuotoiseksi, jolloin ohjelmistot osaavat käsitellä sen. Koneoppiminen mahdollistaa koneen oppimisen ilman, että



ihminen opettaa sitä. Tätä kehittyneemmällä tekoälyllä tarkoitetaan teknologiaa, joka lähestyy inhimillisen älykkyyden tasoa ja jolla on jonkin tasoinen tietoisuus. (Kaarlejärvi & Salminen, 2018, s.51–52). Teknologiasta käytetään yleisesti käytettyä lyhennettä AI (artificial intelligence, suom. tekoäly).



Kuvio 3. Robotiikka ja tekoäly taloushallinnossa. (Kaarlejärvi & Salminen, 2018, s.52)

Tekoälyn sekä ohjelmistorobotiikan käyttöönotto taloushallinnon automaatioasteen nostamiseksi on yksi taloushallinnon digitalisaation voimakkaimpia muutoksia tällä hetkellä (Kaarlejärvi & Salminen, 2018, s.51). Nykyaikaiset tekoälyratkaisut ovat toistaiseksi yhtä käyttötarkoitusta varten rakennettuja, eivätkä yleiskäyttöisiä. Itsenäisen älykkyyden kehittyessä tekoälysovellusten yleiskäyttöisyys lisääntyy, jolloin taloushallinnon tehtävistä on mahdollista automatisoida yli 95 %. Taloushallinnon automatisointiin riittävä automaatio on jo olemassa, koska käsiteltävänä on pääosin digitaalisia aineistoja ja dataa, jotka perustuvat loogisiin sääntöihin ja ovat matemaattisesti mallinnettavissa. (Kaarlejärvi & Salminen, 2018, s.61). Teknologiaa käyttöönottaessa organisaation tulee kuitenkin laskea automatisoinnin kustannukset suhteessa saatuun

hyötyyn ennen tekoälyjärjestelmän hankkimista (Bonnier Pro, 2024, luku 7.1, osa 2.3).

Taloushallinnossa eniten resursseja vievä prosessi on usein ostolaskujen käsittely. Laskujen tarkastus, hyväksyntä ja täsmäytys työllistää myös muuta organisaatiota. (Kaarlejärvi & Salminen, 2018, s.96). Yleensä organisaation koko lisää laskujen käsittelyyn liittyvää byrokratiaa ja useille seurantakohtetasoille kohdistaminen sekä kierrättäminen lisäävät käsittelykustannuksia (Kaarlejärvi & Salminen, 2018, s.103). Manuaalisesti tiliöidyn ja reititetyn ostolaskun kustannus on noin 10 € / ostolasku. Tekoälyä hyödyntämällä voidaan päästä murto-osaan manuaalityön kustannuksista. (Snowfox.AI, 2024). Palkeilla 2022 toteutunut ostolaskun kustannus vuonna 2022 oli 4,22 € / ostolasku (Palkeet 2024d, s.11).

Nykyisissä ostolaskujen käsittelyyn tarkoitetuissa järjestelmissä on mahdollisuus ostotilaukseen tai ostosopimukseen perustuvien laskujen käsittelyyn. Tehokkaimmillaan tilaukselliset tai sopimukselliset laskut eivät vaadi laskun saapuaessa manuaalisia toimenpiteitä, koska hankinnan hyväksymis- ja tiliöintitapahtumat on tehty jo tilaus- ja vastaanottovaiheen aikana (Kaarlejärvi & Salminen, 2018, s. 100). Tilauksellisten ja sopimuksellisten laskujen on täsmätävä ennalta tehtyyn tilaukseen tai sopimukseen saavuttaakseen automaattisen käsittelyn. Edellytyksenä on, että sähköisen laskun sisältämä sanoma sisältää laskun kohdistustiedot oikeassa muodossa. On mahdollista, että tilaukselliseksi tai sopimukselliseksi tarkoitettu lasku ei erinäisistä syistä, kuten tilausnumeron kirjoitusvirheen takia kohdistu automaattisesti tilaukseen, jolloin laskut siirtyvät manuaaliseen käsittelyyn. Kaikista hankinnoista ei myöskään ole mahdollista tai tarkoituksenmukaista tehdä sopimusta tai ostotilausta, jolloin laskun käsittely tiliöintineen ei tapahdu automaattisesti.

Ostolaskun manuaalinen tiliointi on aikaa vievä työvaihe ostolaskujen käsittelyssä. Hitaimmillaan manuaalinen tiliointi tarkoittaa sitä, että laskun käsittelijä tulkitsee laskun kovalta ilman automaation apua, minkälaisesta tuotteesta tai palvelusta on kyse ja miten se tulisi kirjanpidossa käsitellä. Laskun

manuaalinen käsittely voi tällöin vaatia tilikartan tutkimista ja päättelyä, miten lasku tulisi tiliöidä.

Tiliöinnillä tarkoitetaan kirjanpidon tapahtuman kohdistamista oikealle kirjanpidon riville, että se voidaan käsitellä oikein (Riihos, 2024). Tiliöinnin merkitys kirjanpidossa on tärkeä, koska se mahdollistaa kirjanpidon tapahtumien seurannan, analysoinnin ja raportoinnin. Tiliöinnin avulla ylläpidetään tarkkoja kirjanpidon tietoja ja varmistetaan, että jokainen tapahtuma kirjataan oikein ja oikealle tilille. (Briox, 2024). Yleiset kirjanpitoperiaatteet ja esimerkiksi tilinpäätöslaskelmien muoto ovat yhtäläisiä niin julkisella kuin yksityisellä sektorilla. Erot kirjanpidon fokuksessa näkyvät siten, että yksityisellä sektorilla seurataan organisaation tulosta ja sen muodostumista, kun taas julkisella sektorilla fokus on menojen lainmukaisuuden ja talousarvion toteutumisen seurannassa. (Palkeet 2024c).

Palkeilla sekä valtion taloushallinnossa yleisesti tiliöinti perustuu Valtiokonttorin määräämään tilikarttaan. Valtiokonttori antaa määräyksen muun muassa liikekirjanpidon tilikartasta, talousarviokirjanpidon tileistä ja siirrettyjen määrärahojen tilien muodostamisesta (Valtiokonttori, 2024). Lakisäänteisten kirjanpidon tilien lisäksi Palkeilla on tiliöitävänä usein esimerkiksi virastokohtaisia seurantakohteita, joiden avulla asiakasvirastoille on mahdollista seurata yksityiskohtaisempia kirjanpidon tapahtumia, kuten jonkin viraston sisäisen projektin menoja.

Palkeilla yksittäinen ostolasku voi saada kirjaukset liikekirjanpidon ja talousarviokirjanpidon tileille sekä useille seurantakohteille, joka tekee tiliöintiriveistä pitkiä ja monimutkaisia. Tiliöintirivien täydentäminen sekä tarkastaminen manuaalisesti on paljon resursseja vaativa työvaihe. Haaste on tunnistettu sekä yksityisellä että julkisella puolella ja avuksi on kehitetty tekoälyyn ja koneoppimiseen perustuvia ratkaisuja. Palkeilla aihe on ollut ajankohtainen, kun Snowfox.AI-ostolaskutekoäly otettiin käyttöön osaksi ostolaskujen tiliöintipalvelua keväällä 2022. Kyseessä on tekoälyratkaisu, joka perustuu koneoppimiseen. Snowfox.AI ei korvaa ihmistä, mutta nopeuttaa ja tehostaa laskujen käsittelyä

luomalla tiliöintiennusteet lasku- ja tiliöintihistoriadataan perusteella. (Palkeet, 2024a).

Historiatietojen ja olemassa olevan datan hyödyntäminen on olennaista koneoppivan tekoälyn kouluttamisessa. Palkeiden Snowfox.AI sai ennen käyttöönottoa Palkeilla käsiteltyjen laskujen kuvat (PDF), sanomat (XML) sekä niiden toteutuneet tiliöinnit yhdeksän kuukauden ajalta. Tekoälyllä on käytössä myös valtion yhteiset ja asiakaskohtaiset validointisäännöt sekä seurantakohteet. (Palkeet 2024a). Koulutukseen tarvittava data täytyy valmistella ja luoda datasta yhtenäinen ja johdonmukainen tietokanta. Valmisteltu data jaetaan kolmeen eri ryhmään, koulutus-, testaus- ja validointidataan. Testaus- ja validointidatalla varmistetaan, että algoritmi toimii hyvin myös datalla, jota se ei ole koskaan nähnyt. (Kananen & Puolitaival, 2019, s.46). Snowfox.AI:n koulutuksessa käytetään 70 % toimitetusta datasta koulutukseen ja 30 % testaukseen ja validointiin (Snowfox.AI, 2024).

Snowfox.AI:n tarjoaa tiliöntiehdotuksen historiadataan tuottamien ennusteiden perusteella. Ennusteet pohjautuvat siihen, minkälaisia tiliöintejä vastaaville ostolaskuille on tehty aikaisemmin. Palvelu tarkistaa tekemänsä ennusteet validointisääntöjä vasten ja tarvittaessa korjaa tai täydentää ennusteet. Ennusteen luotettavuuden arvioimiseksi palvelu vie ennusteelle luottamusarvon ennusteen tarkkuudesta prosentteina (Palkeet, 2024a).

Snowfox.AI vie validointisääntöjen mukaiset tiliöintiennusteet laskun XML-sanomalle ja Palkeiden ostolaskujen käsittelyjärjestelmä Handin laskukohtaisille tiliöintiriveille. Laskun käsittelijä tarkastaa ja tarvittaessa korjaa tekoälyn tekemät tiliöintiennusteet. (Palkeet, 2024a). Koneoppimisen avulla ei aina saada sataprosenttisesti oikeaa tiliöintiä, mutta jos sen avulla saadaan tiliöintitiedot osittain oikein ihmisen käsittelyä varten, tehostaa se käsittelyä verrattuna manuaaliseen työhön (Kananen & Puolitaival, 2019, s.105).

Palkeilla tekoälyn onnistumisaste vuonna 2023 oli noin 36 %. Parannusta edellisvuoteen oli noin kolme prosenttiyksikköä. (Palkeet 2024b). Vielä on matkaa valmistajan lupaamaan yli 90 % automaatioasteeseen (Snowfox.AI, 2024).

Ennustetarkkuuden jääminen reiluun kolmasosaan valmistajan lupaamasta selittynee osittain julkishallinnon organisaatioiden ostolaskujen pitkistä tiliointiriveistä, joissa on yksityisen sektorin organisaatioiden ostolaskuihin verrattuna oletetusti enemmän seurantakohteita sekä talousarviokirjanpidon tilejä, jotka monimutkaistavat ennusteiden tekemistä.

Snowfox.AI:n onnistumisaste kertoo tekoälyn oikein tiliöimien ostolaskujen osuuden, joka on toki hyödyllistä palvelun toiminnan tason hahmottamisessa. Onnistumisasteesta ei voida kuitenkaan vetää suoria johtopäätöksiä tekoälyn vaikutuksesta ostolaskujen käsittelyn tuottavuuteen. Tuottavuuteen vaikuttavia tekijöitä on lähes rajattomasti - vaikuttavia tekijöitä voi olla esimerkiksi järjestelmien käytettävyys, prosessien tehokkuus ja se, saadaanko toimittajilta ostolaskut sellaisessa muodossa, että mahdollisimman tehokas automaatio olisi ylipäättään mahdollista.

### 3.3 Tuottavuus julkishallinnossa

Tuottavuus voidaan määritellä organisaatiossa aikaansaadun tuotannon määrän ja tuotannontekijöiden käytön väliseksi tuotos-panos-suhteeksi (Neilimo & Uusi-Rauva, 2017, s.281). Panos voi olla henkilöstöä tai muuta pääomaa, kuten koneita tai laitteita. Organisaation tuottavuuden lähtökohtia ovat henkilöstön työpanos, viihtyvyys, taidot ja osaaminen, käytössä oleva teknologia ja laitteisto sekä johdon taito johtaa näistä muodostuvaa kokonaisuutta. (Neilimo & Uusi-Rauva, 2017, s.312). Varsinkin yksityisellä sektorilla tuottavuus liitetään usein kannattavuuteen ja tehokkuuteen ja tuottavuuden parantaminen yhdistetään usein voiton tavoitteluun. (Herrmann & Tarren, 2012, s. 2).

Voiton tavoittelu ei ole moitittavaa myöskään julkisella sektorilla, muttei yleensä päämäärä. Julkisella sektorilla huomioonotettavia asioita ovat muun muassa laatu, saatavuus, vastuunalaisuus sekä kustannusten hallinta (Herrmann & Tarren, 2012, s. 2). Julkisen sektorin organisaatioiden tavoite kytkeytyy vaikuttavuuden aikaansaamiseen ilman voiton tavoittelua, joskin useimmiten vahvasti talousohjattuna (Listenmaa, 2023, s. 78). Tuottavuustavoitteet

ovat tärkeä osa organisaation kannattavuustavoitteen saavuttamisessa. Hyvä tuottavuus mahdollistaa tehokkaan kustannustehokkuusstrategian käytön ja parantaa siten kustannuskilpailukykyä. (Neilimo & Uusi-Rauva, 2017, s.282). Palkeiden näkökulmasta tuottavuuden parantamisella pyritään kustannustehokkuuteen, joka on yksi Palkeiden perustehtävistä.

Tuottavuuden mittaamiseen liittyy omat haasteensa, kuten miten mitattava tuotos tai tulos määritellään. Prosessi on suhteellisen suoraviivainen, mikäli tuotantoprosessin tuotoksena syntyy konkreettisia, laskettavissa olevia hyödykkeitä, kuten esimerkiksi autotehtaan tuottamien autojen lukumäärä viikossa. (Herrmann & Tarren, 2012, s. 2). Rajasin opinnäytetyöni aiheen koskemaan ostolaskujen käsittelyä, koska ostolaskujen käsittelyn tuotoksena syntyy selkeitä suoritemääriä, eli reskontraan/maksuun saakka saatuja ostolaskuja. Esimerkiksi kaikkien Palkeiden talouspalveluiden tuottavuuden mittaaminen olisi haasteellista palveluiden erilaisten prosessien takia, koska eri prosesseista olisi mahdotonta löytää yhteismitallisia suoritteita.

Ostolaskujen käsittelyn tuottavuutta mitataan opinnäytetyöni tuottavuusaineistossa muodossa tehollinen henkilötyövuosi (HTV) / suoritekirjausten määrä. Henkilötyövuosi kuvaa kokoaikaiseksi muutetun henkilön työpanosta, joten yhteisön kaikki palkatut tunnit jaetaan yrityksen kokoaikaisten palkansaajien keskimääräisillä palkatuilla tunneilla vuodessa (Tilastokeskus, 2024a). Tuottavuusaineiston tehollinen HTV muodostuu ostolaskujen käsittelylle kirjattujen tuntien perusteella muodossa kirjattut työtunnit / vuosittaisten työpäivien määrä / päivittäisten työtuntien määrä (=työtunnit / 253 / 7,25). Lineaarisen regressioanalyysin avulla selvitän, kuinka suuren osuuden sataprosenttisesti onnistunut AI-tiliöinti selittää tuottavuudesta.

## 4 REGRESSIOMALLIN MUODOSTAMINEN JA TULOKSET

### 4.1 Lineaarisen regressioanalyysin perusteet

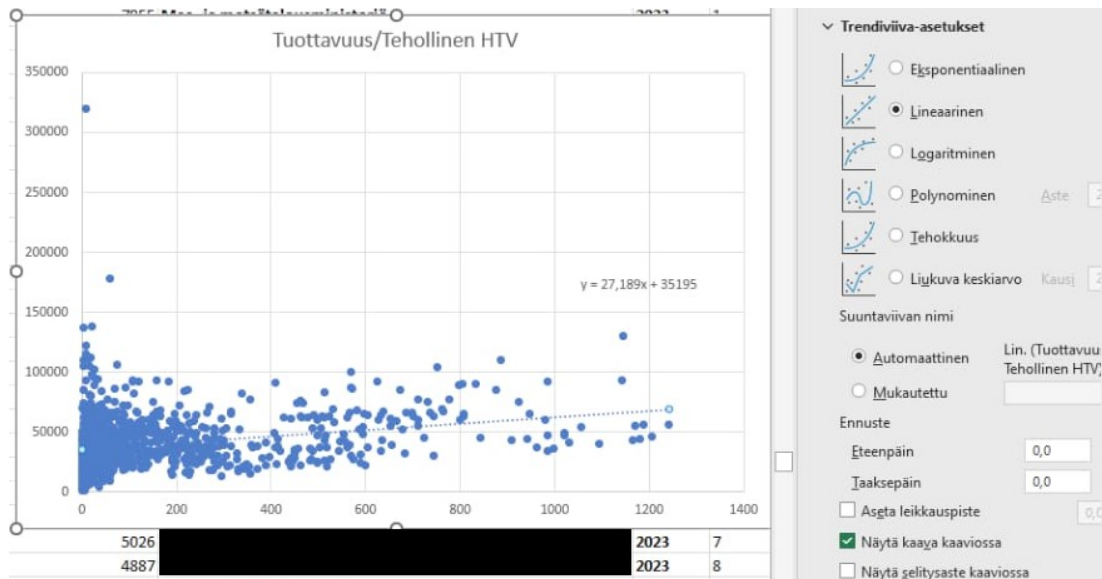
Työn regressioanalyysi on tehty Microsoft Exceliin saatavilla olevan analyysityökalupaketin regressio-lisätyökalulla. Ennen tietotekniikan kehitystä tilastollinen tutkimus oli pääasiassa laskemista. Ohjelmistojen kehittyttyä tilastollisessa tutkimuksessa ei nykyisin juuri vaadita käsin laskemista, koska kone tekee sen tutkijan puolesta. On kuitenkin tärkeää ymmärtää mistä ohjelmiston, kuten Excelin laskemat tulokset muodostuvat, jotta tuloksia voidaan tulkita oikein. (Valli, 2015, s.10).

Regressiomalleja voidaan muodostaa eri tavoilla, mutta yleisin malli on regressiosuora, joka mallintaa kahden muuttujan välistä lineaarista yhteyttä. Malli on käyttökelpoinen, koska useat luonnon, ihmismielen ja talouden ilmiöt voidaan mallintaa melko tarkasti lineaarisen mallin avulla. Regressiomallista käytetään merkintätapaa

$$\hat{y} = b_0 + b_1 \cdot x, \quad (1)$$

jossa  $\hat{y}$  = Y-muuttujan (selitettävän muuttujan) ennustettu arvo,  $b_0$  = regressiosuoran vakiotermin,  $b_1$  = regressiokerroin ja  $x$  = X-muuttujan (selittävän muuttujan) arvo. (Nummenmaa et al., 2019, s.237).

Excelissä kaava saadaan muodostettua esimerkiksi muodostamalla pistekaavio selitettävän ja selittävän muuttujan arvoista. Kaavioon voidaan lisätä regressiosuora sekä kaava lisäämällä kaavion osa -> suuntaviiva (lineaarinen) ja muotoilemalla suuntaviivaa -> näytä kaava kaaviossa (Kuvio 4). Kaavaa voidaan käyttää tuottavuuden ennustamiseen; lisäämällä sataprosenttisesti oikein AI-tiliöityjen laskujen suoritettua x:n paikalle saadaan y:n arvoksi ennustettu tuottavuus (kpl / tehollinen HTV).



Kuvio 4. Kuvaruutukaappaus Excelin pistekaaviosta, regressiosuorasta ja kaavasta.

Regressiomallinnuksen varsinainen ongelma on vakiotermin  $b_0$  sekä regressiokertoimen  $b_1$  määrittäminen. Matemaattisesti yksinkertaisin tapa on pienimmän neliösumman menetelmä, joka perustuu regressioanalyysin jäännöstermien tarkastelemiseen (Nummenmaa et al., 2019, s.238). Excel laskee vakiotermin  $b_0$  sekä regressiokertoimen  $b_1$  automaattisesti, mutta aiheen ymmärrettävyyden takia on aiheellista käydä läpi teoria pääpiirteittäin.

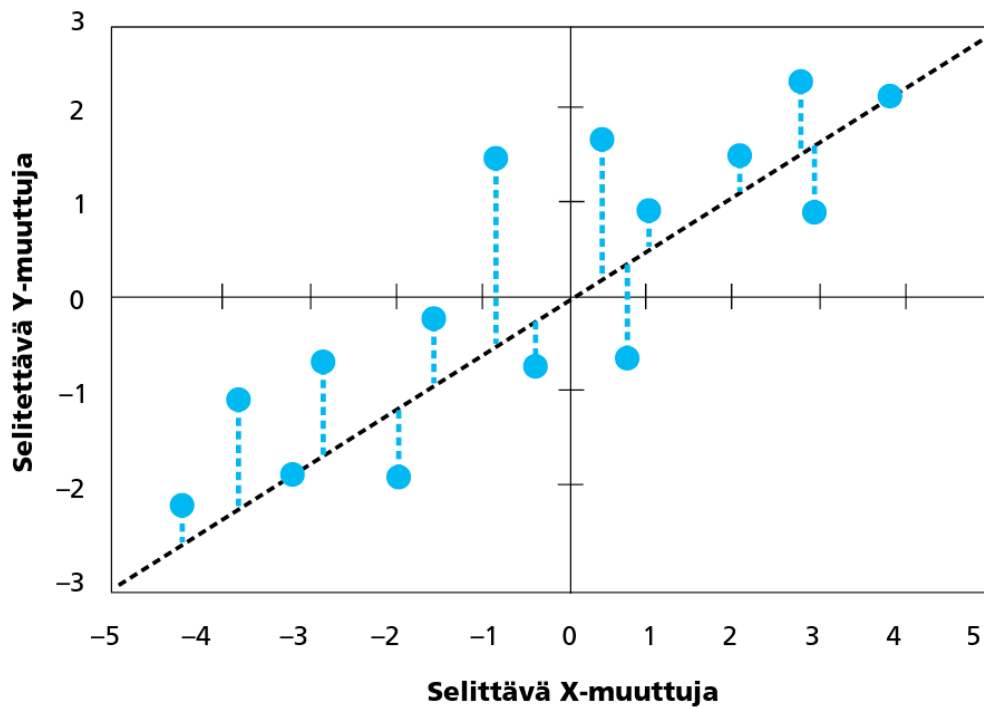
Regressiomallia käyttäessä Y-muuttujalle voidaan ennustaa arvo  $\hat{y}$ . Ennuste ei kuitenkaan ole täydellinen ja se voi poiketa huomattavasti todellisesta Y-muuttujan arvosta. Näitä poikkeamia  $(y_i - \hat{y}_i)$  kutsutaan jäännöstermeiksi ja merkinnäksi  $e$  (sanasta error). Pienimmän neliösumman menetelmällä pyritään määrittämään sellainen regressiosuora, jota käytettäessä neliöityjen jäännöstermien yhteenlaskettu summa

$$\text{Min} \sum (y_i - \hat{y}_i)^2 = \text{Min} \sum e_i^2 \quad (2)$$

olisi mahdollisimman pieni. Jäännöstermit neliöidään, etteivät positiiviset ja negatiiviset termit kumoaisi toisiaan summalausekkeessa. Kun neliösumma on niin pieni kuin mahdollista, kuvaa regressiosuora aineiston mahdollisimman hyvin. (Nummenmaa et al., 2019, s.238).



Kuviossa 5 pistejoukkoa mallinnetaan sen halki kulkevalla regressiosuoralla. Pystysuorat katkoviivat kuvaavat jäännöstermien etäisyyttä regressiosuoraan. Eri asennoissa regressiosuora jäännöstermien neliöiden summa pienenee tai kasvaa. Mitä pienempi neliösumma on, sitä paremmin regressiosuora ”sopii” aineistoon. (Nummenmaa et al., 2019, s.238).



Kuvio 5. Kuva havainnollistaa pienimmän neliösumman menetelmää (Nummenmaa et al., 2019, s.238).

Excel muodostaa regressiosuoran pienimmän neliösumman menetelmällä sekä laskee samalla vakiotermin  $b_0$  sekä regressiokertoimen  $b_1$ . Vakiotermi  $b_0$  ilmoittaa pisteen, jossa suora leikkaa Y-akselin, eli Y-muuttujan arvon silloin, kun X-muuttujan arvo on nolla. Termi  $b_1$  ilmaisee suoran kulmakertoimen, eli kuinka paljon ja mihin suuntaan Y:n arvo muuttuu, kun x kasvaa yhden yksikön verran. (Nummenmaa et al., 2019, s.239). Excelin muodostaessa parametrien arvot automaattisesti ei parametrien  $b_0$  ja  $b_1$  manuaalinen laskeminen ole opinäytetyöni kannalta olennaista, joten todettakoon näille löytyvän valmiit kaavat.

$$b_1 = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{n(\sum x_i^2) - (\sum x_i)^2} \quad (3)$$

$$b_0 = \frac{\sum y_i - b_1(\sum x_i)}{n} = \bar{y} - b_1 \bar{x} \quad (4)$$

## 4.2 Tutkimusaineiston koostaminen

Tutkimuksessani selitettävä muuttuja on ostolaskujen käsittelyn tuottavuus (kpl / tehollinen HTV) ja selittävä muuttuja Snowfox.AI:n sataprosenttisesti oikein tiliöimien ostolaskujen määrät. Tutkimusaineiston koostaminen alkoi järjestelmistä muodostettujen aineistojen järjestelyllä analysoitavaan muotoon. Tutkimusaineisto koottiin Excel-taulukoksi, jossa myös lineaarinen regressioanalyysi toteutettiin.

Snowfox.AI:n osalta käytössäni oli kuukausittaiset raportit palvelun käyttöönotosta työn empiriaosan toteutukseen saakka, yhteensä 19 kuukauden ajalta, 04/2022–10/2023. Aineisto sisälsi kuukausittaiset virastokohtaiset Snowfox.AI:n ennustamien laskujen suoritemäärät sekä tiliöintiennusteiden onnistumisprosentit eriteltyinä kolmella tavalla:

1. kaikki onnistuneet tiliöinnit;
2. kaikki onnistuneet yksiriviset tiliöinnit;
3. kaikki onnistuneet usean rivin tiliöinnit.

Tuottavuuden osalta käytössäni oli aineisto, johon oli koottuna asiakaskohtaiset kuukausittaiset suoritemäärät eriteltyinä eri ostolaskutyypeille. Järjestelmä oli laskenut valmiiksi suoritemäärien sekä eri ostolaskutyypeille kirjattujen työaikojen perusteella tehollisen HTV:n sekä tuottavuuden. Tuottavuuteen liittyvää aineistoa käytettävissäni oli vuoden 2022 alusta työn empiriaosuuden toteutukseen saakka, yhteensä 22 kuukauden ajalta (01/2022–10/2023).

### 4.3 Tutkimusaineiston yhdistely

Lineaarisen regressioanalyysin toteutus Microsoft Excelissä vaatii aineiston järjestelyn siten, että selittävät ja selitettävät muuttujat ovat yhtenäisessä ja vertailtavassa muodossa samassa taulukossa. Snowfox.AI:n onnistuneiden tiliöintiennusteiden määrät ja tuottavuudet tuli saada samoille kuukausittaisille asiakaskohtaisille riveille. Snowfox.AI:n osalta tutkimusaineisto oli selkeää, koska jokaiselta tarkasteltavalta kuukaudelta löytyi yksi asiakaskohtainen rivi, joka sisälsi tutkimuksen kannalta tarvittavan tiedon (Kuvio 6).

Kirjanpitoyksikkö	Laskuja ennustettu	Onnistumisprosentti	Onnistuneet kpl	Päiväys
Asiakas A	67	28,36 %	19	30.7.2022
Asiakas B	1096	32,39 %	355	30.8.2022
Asiakas C	2112	36,79 %	777	30.6.2023
Asiakas D	3512	20,27 %	712	30.12.2022
Asiakas E	210	25,24 %	53	30.4.2022
Asiakas F	1204	38,29 %	461	31.7.2023
Asiakas G	1935	37,83 %	732	30.8.2022
Asiakas H	409	36,92 %	151	30.10.2022

Kuvio 6. Kuva SnowfoxAI:n tiliöinnin ennusteista sekä onnistuneista tiliöinneistä.

AI-tiliöinnin osalta lähdeaineistoa oli noin 1 200 riviä ja tuottavuuksien osalta noin 3 800 riviä, joten päädyin tiivistämään tuottavuusaineiston rivimäärältään vastaavaan muotoon AI-aineiston kanssa. Tuottavuusaineisto oli laajempi, koska tuottavuudet oli laskettu eri ostolaskutyypeille. Tuottavuuksien osalta aineistoa oli myös pidemmältä ajanjaksolta. Tarkasteltaessa AI-tiliöinnin vaikutusta ostolaskujen käsittelyn kokonaistuottavuuteen, ei ole syytä eritellä eri ostolaskutyyppejä omille riveilleen (Kuvio 7). Eri ostolaskutyyppeirivien yhdistäminen oli siis mahdollista, jolloin aineisto saatiin samaan muotoon AI-aineiston kanssa.

Kirjanpitoyksikkö	Ostolaskutyyppi	Suoritekirjaukset	Tehollinen HTV	Tuottavuus / Tehollinen HTV	Vuosi	Kuukausi
Asiakas A	Muut ostolaskut	76	0,0017	44 706	2023	10
Asiakas B	Osittain automaattiset ostolaskut	849	0,0190	44 684	2022	9
Asiakas C	Manuaaliset ostolaskut	89	0,0020	44 500	2022	9
Asiakas D	Manuaaliset ostolaskut	31	0,0007	44 286	2022	7
Asiakas E	Muut ostolaskut	2 345	0,0531	44 162	2023	3
Asiakas F	Manuaaliset ostolaskut	3 460	0,0785	44 076	2022	5
Asiakas G	Muut ostolaskut	254	0,0058	43 793	2023	6
Asiakas H	Manuaaliset ostolaskut	381	0,0087	43 793	2022	12

Kuvio 7. Kuva tuottavuusaineistosta. Asiakaskohtaiset suoritekirjaukset, tehollinen HTV sekä tuottavuus eriteltynä eri ostolaskutyypeille.

Tuottavuusaineistoa jouduttiin täydentämään työn edetessä, koska ostolaskutyyppien kategorisointi oli muuttunut vuosien 2022 ja 2023 välillä ja osa aineistosta oli tästä syystä jäänyt ensimmäisestä aineistosta pois. Aineisto vaati lisäksi tarkasteltavan ajanjakson ulkopuolelle jäävien rivien poistamisen. Tämän jälkeen aineiston yhdistely onnistui Excelin Pivot-taulukoinnin avulla, jolla kuukausittaiset asiakaskohtaiset suoritelmäärät, tehollinen HTV sekä tuottavuus saatiin yhdistettyä, suodatettua ja järjestettyä samaan muotoon AI-aineiston kanssa (Kuvio 8).

Kirjanpitoyksikkö	Suoritekirjaukset	Tehollinen HTV	Tuottavuus/Tehollinen HTV	Vuosi	Kuukausi
Asiakas A	908	0,0114	79649	2022	7
Asiakas B	3019	0,039	77410	2022	8
Asiakas C	3830	0,0497	77062	2023	6
Asiakas D	9367	0,1217	76968	2022	12
Asiakas E	315	0,0041	76829	2022	4
Asiakas F	3015	0,0399	75564	2023	7
Asiakas G	3149	0,042	74976	2022	8
Asiakas H	772	0,0103	74951	2022	10

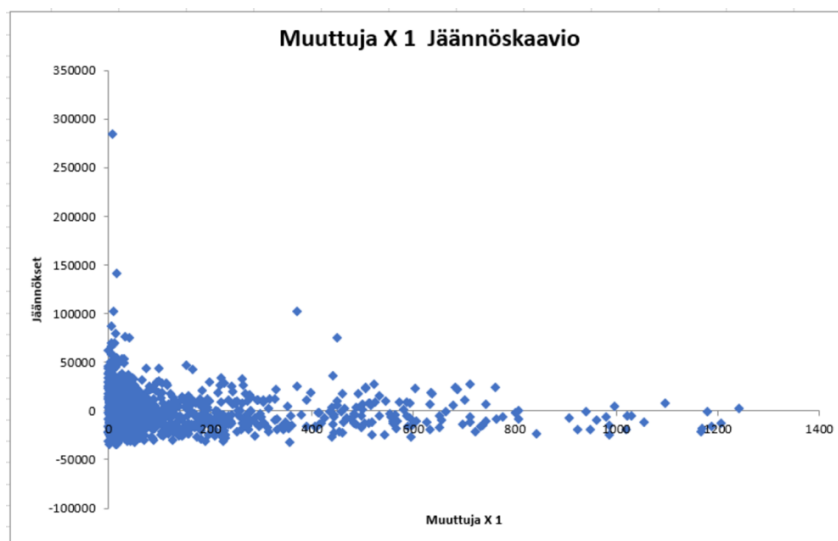
Kuvio 8. Kuva tuottavuusaineistosta. Tuottavuudet yhdistetty ja pivotoitu samaan muotoon AI-raportin kanssa.

Aineiston yhdistämisen jälkeen tarkistin, että yhdistetyt suoritelmäärät, HTV sekä tuottavuus täsmäsivät alkuperäisiin järjestelmistä muodostettuihin lukuihin. Tietojen pitäessä paikkansa aineisto voitiin yhdistää AI-aineiston kanssa samalle sivulle Excelissä. Aineistojen ollessa samalla sivulla sitä voitiin suodattaa ja järjestellä halutulla tavalla samassa näkymässä.

#### 4.4 Poikkeavat havainnot ja aineiston muokkaus

Aineiston ollessa järjestetty haluttuun muotoon, ajoin Excelissä regressioanalyysin sekä jäännöskuvaajan tarkasteltavalta ajanjaksolta (04/22–10/23).

Jäännöskuviaita kannattaa silmäillä mahdollisten poikkeavien havaintojen varalta. Yksittäiset poikkeavat pisteet viittaavat poikkeaviin havaintoihin, joiden poisjättämistä mallista on mietittävä. (Taanila, 2020, s.14). Joskus poikkeavien havaintojen taustalla voi olla yksinkertainen koodausvirhe tai virhe aineistossa, mutta useimmiten kyse on siitä, että jotkin havainnot saavat todellisuudessa muista poikkeavia arvoja. Tällaisessa tilanteessa tulee arvioida, löytyykö poikkeaville havainnoille mitattavissa oleva selitys, jolloin havainto voidaan sisällyttää aineistoon uutena muuttujana eikä se silloin vääristä tutkimustulosta. Joissain tapauksissa poikkeavat havainnot voidaan perustellusti jättää pois aineistosta, mikäli ne vääristävät tutkimustuloksia (Tietoarkisto, 2024c). Huomioni kiinnittyi jäännöskaavion muutamisiin yksittäisiin poikkeaviin havaintoihin muuten siistissä jäännöskuviossa (Kuvio 9).

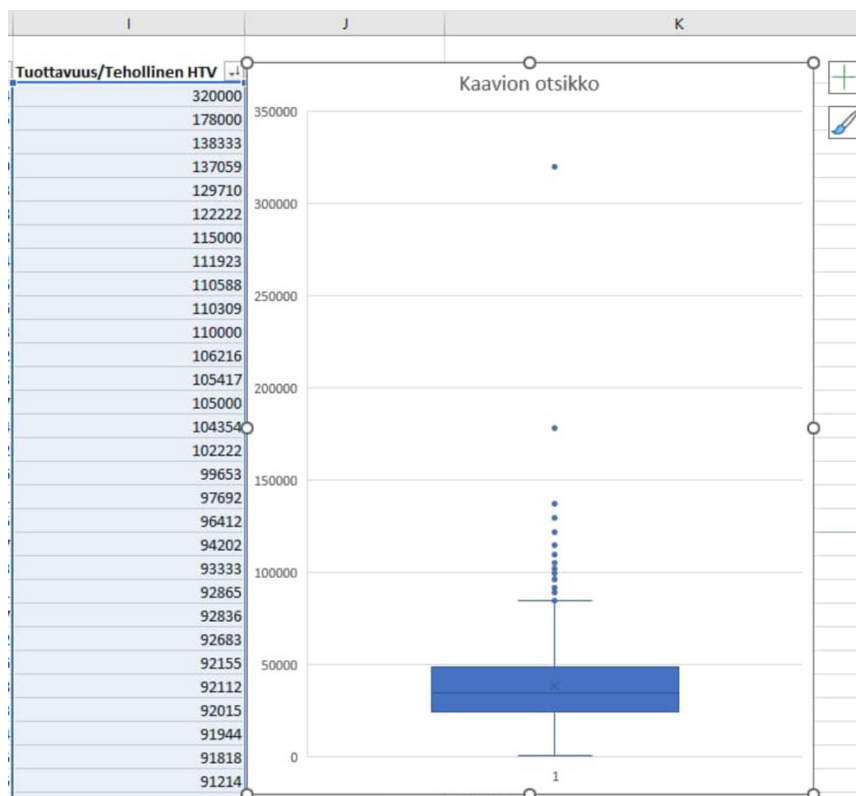


Kuvio 9. Poikkeavat havainnot erottuvat yksittäisinä pisteinä kaavion yläosassa ensimmäisen regressioajon jäännöskaaviossa.

Kuvaajassa näkyvät poikkeavat havainnot viittasivat tuottavuusaineiston riiveille, joilla tuottavuuslukemat näyttivät poikkeuksellisen korkeita lukemia. Palkeiden taloushallinnon erityisasiantuntijan Niko Heikkisen kanssa asiaa pohdittaessa päätimme, että poikkeavat havainnot johtuivat todennäköisesti virheellisistä tai epätarkoista työajan kirjauksista. Tällainen poikkeama voi syntyä esimerkiksi ajanjaksona, jolloin asiakasvirastolla on vähän ostolaskuja, jolloin myös niiden käsittelyyn käytetty työaika on vähäistä. Vähäisillä työtunneilla

työajankirjaukset voivat olla puutteellisia tai jopa jäädä tekemättä, jolloin kyseisen ajanjakson tuottavuuslukemat näyttävät poikkeuksellisen suurilta.

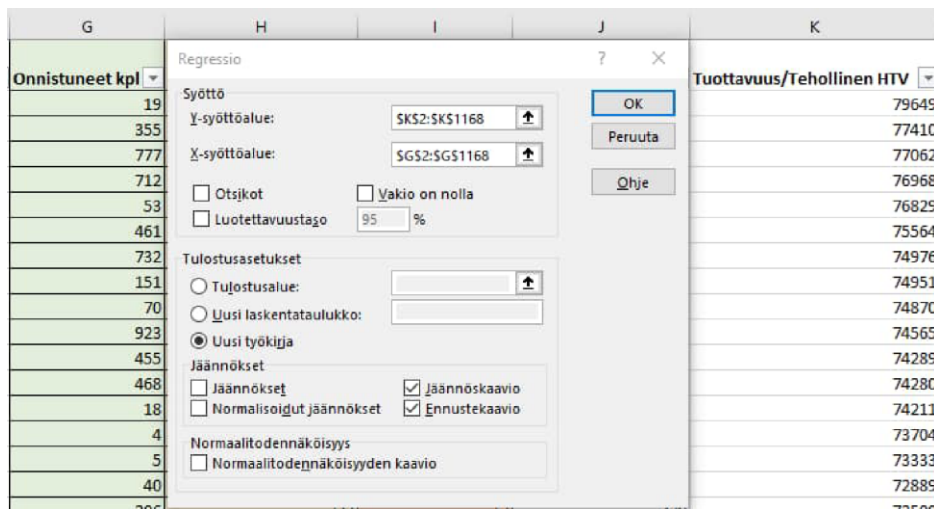
Päädyin poistamaan aineistosta poikkeuksellisen suurta tuottavuutta osoittavat rivit, koska pidin niiden tuottavuuslukuja mahdottomina ja siten vääristävän tuloksia. Suodatin aineiston tuottavuuksien mukaan suurimmasta pienimpään ja hyödynsin Excelin ”ruutu ja jana”-kaaviota (Kuvio 10). Kuvaajan sininen ruutu sisältää 50 % jakauman keskiosan tapauksista. Ruudun pohja on alakvartiili, kansi yläkvartiili ja ruudun poikki menevä jana mediaani (Heikkilä, 2014, s.166). Ruudun pystysuuntaiset janat osoittavat vaihtelua ylä- ja alaneljännesten ulkopuolella ja näiden ulkopuolella olevia pisteitä pidetään poikkeavina (Microsoft, 2023). Poistin aineistosta suurinta tuottavuutta osoittavat rivit, kunnes kuvaajan poikkeavat havainnot saatiin pois. Tämän jälkeen aineisto oli valmis varsinaista analyysiä varten.



Kuvio 10. Excelin ”ruutu ja jana”-kuvaaja erottaa poikkeavat havainnot pisteinä.

#### 4.5 Regressioanalyysin tulokset

Regressioanalyysin muodostaminen Excelissä on yksinkertaista, kun aineisto on valmisteltu taulukkoon analysoitavaan muotoon siten, että selittävän ja selitettävän muuttujan tiedot ovat samoilla riveillä. Excelin regressiotyökaluun määritellään Y-syöttöalueeksi selitettävän muuttujan tiedot ja X-syöttöalueeseen selittävän muuttujan tiedot (Kuvio 11). Työkalulla saa samalla muodostettua ennuste- ja jäännöskaaviot analyysistä.



Kuvio 11. Excelin regressiotyökalu.

Excel muodostaa aineistosta Regressiotunnusluvut-taulukon. Taulukon kerroin R tarkoittaa korrelaatiokerrointa. Excelin suomenkielisessä versiossa selityskerrointa kutsutaan virheellisesti korrelaatiokertoimeksi, eli taulukon korrelaatiokerroin tarkoittaa selityskerrointa (Taanila, 2020, s.8). Selkeyden vuoksi nimet on korjattu alla olevaan taulukkoon 2. Korrelaatiokerroin mittaa lineaarisen yhteyden voimakkuutta selittävän ja selitettävän muuttujan välillä. Korrelaatiokerroin voi saada arvoja välillä -1 ja 1. Arvo 1 kertoo täydellisestä positiivisesta yhteydestä muuttujien välillä, eli toisen muuttujan arvot kasvavat samassa suhteessa toisen muuttuessa. Arvolla -1 tilanne on päinvastainen. (Tietoarkisto, 2024a). Korrelaatiokerroin 0,2875 osoittaa jonkin asteista, muttei kovin voimakasta positiivista korrelaatiota muuttujien välillä. Voidaan kuitenkin tulkita, että sataprosenttisesti oikein tiliöityjen ostolaskujen suoritemäärillä on positiivinen yhteys tuottavuuteen eli oikein tiliöityjen suoritemäärien kasvaessa myös tuottavuus kasvaa.

Taulukko 2. Excelin Regressiotunnusluvut-taulukko (nimet korjattu).

<i>Regressiotunnusluvut</i>	
Korrelaatiokerroin	0,287528376
Selityskerroin	0,082672567
Tarkistettu selityskerroin	0,081885161
Keskivirhe	15390,68658
Havainnot	1167

Selityskerroin 0,0827 mittaa mallin kykyä kuvata selitettävän muuttujan vaihtelua. Selityskerroin voidaan laskea korrelaatiokertoimen avulla seuraavasti

$$R^2 = r^2 \cdot 100 \% \quad (5)$$

Selityskerroin ilmaisee, kuinka monta prosenttia selitettävän muuttujan Y-arvojen vaihtelusta voidaan selittää selittävän muuttujan X avulla. (Nummenmaa et al., 2019, s.251). Yllä olevan taulukko 2:n mukaan selityskerroin on 0,08267, eli Snowfox.AI:n oikein tiliöimät laskut selittävät tuottavuuden vaihtelusta 8,27 %. Regressiotunnusluvut-taulukosta on luettavissa lisäksi ”tarkistettu selityskerroin”, jota käytetään silloin, kun halutaan verrata kahden regressioanalyysin tuloksia keskenään. ”Keskivirhe” kertoo residuaalien, eli jäännösten keskihajonnan ja ”havainnot” aineiston analysoitujen rivien määrän. (Tietoarkisto, 2024b).

Excelin tulostaulukon kertoimet-sarakkeesta saadaan regressiomallin kertoimet  $b_0$  ja  $b_1$  (Taulukko 3). Vakiotermi  $b_0$  33252 kertoo regressiosuoran Y-akselin leikkauspisteen, eli tuottavuuden tilanteessa, jossa AI ei ennustaisi laskuja ollenkaan.  $b_1$  22,64 ilmaisee suoran kulmakertoimen, eli  $b_1$ :n perusteella voidaan tulkita, että yksi Snowfox.AI:n oikein ennustama ostolasku merkitsee noin 23 laskua/tehollinen HTV lisää tuottavuutta. Mallin tilastollista merkittävyyttä voidaan tulkita taulukko 3:n p-arvosta. Jokaisen tilastollisen tutkimuksen tuloksena saadaan p-arvo, joka ilmoittaa virheellisen päätelmän todennäköisyyden. Yleisesti käytetään 0,050, eli viiden prosentin tai 0,010, eli yhden prosentin riskitasoa. Viiden prosentin riskitaso tarkoittaa, että tulos on tutkimuksen perusjoukossa 95 %:n varmuudella pätevä, mutta virheen

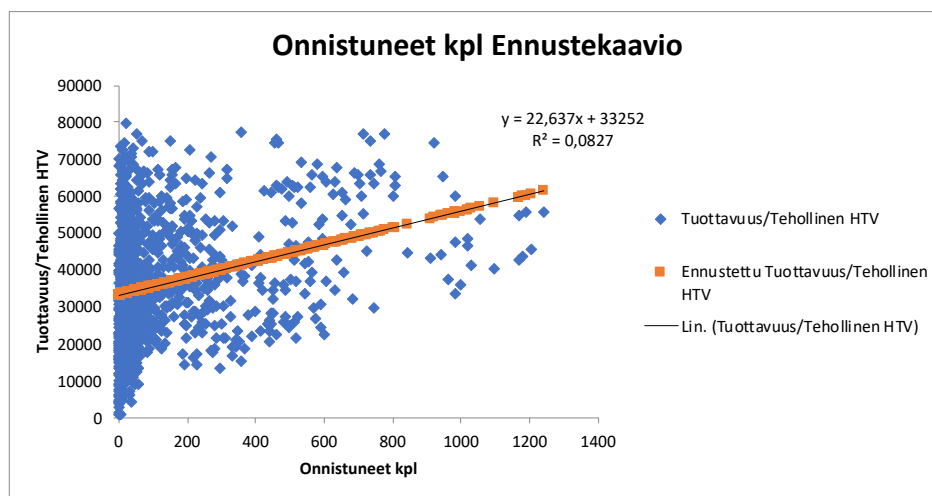


todennäköisyys on 5 %. (Tietoarkisto, 2024d). Mikäli p-arvo on alle 0,050, mallia voidaan pitää tilastollisesti ”melkein merkitsevänä”. Jos p-arvo on alle 0,010, mallia voidaan pitää ”merkitsevänä” ja alle 0,001 mallia voidaan pitää ”erittäin merkitsevänä”. (Tietoarkisto, 2024d; Taanila, 2020, s.15–16). Aineiston p-arvo 1,19239E-23 on alle 0,001, joten mallia voidaan pitää tilastollisesti erittäin merkitsevänä.

Taulukko 3. Excelin tulostaulukko

	<i>Kertoimet</i>	<i>Keskivirhe</i>	<i>t Tunnusluvut</i>	<i>P-arvo</i>
Leikkauspiste	33252,05163	526,7123694	63,13132853	0
Onnistuneet kpl	22,63675405	2,20918797	10,2466401	1,19239E-23

Excelin regressiotyökalulla saa tulostettua lisäksi ennustekaavion mallista (Kuvio 12). Taulukkoon saa sisällytettyä regressiomallin kaavan sekä selityskertoimen. Kuviosta löytyy mallin kaava  $y = 22,637x + 33252$  sekä edellä mainittu selityskerroin 0,0827.

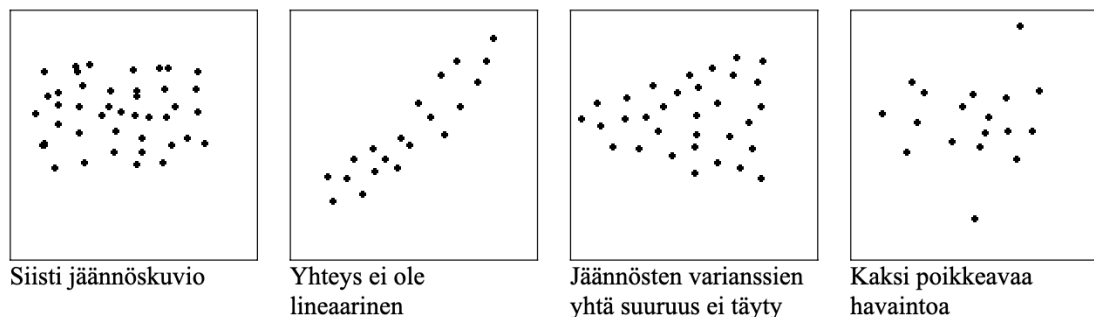


Kuvio 12. Excelin regressioanalyysin ennustekaavio.

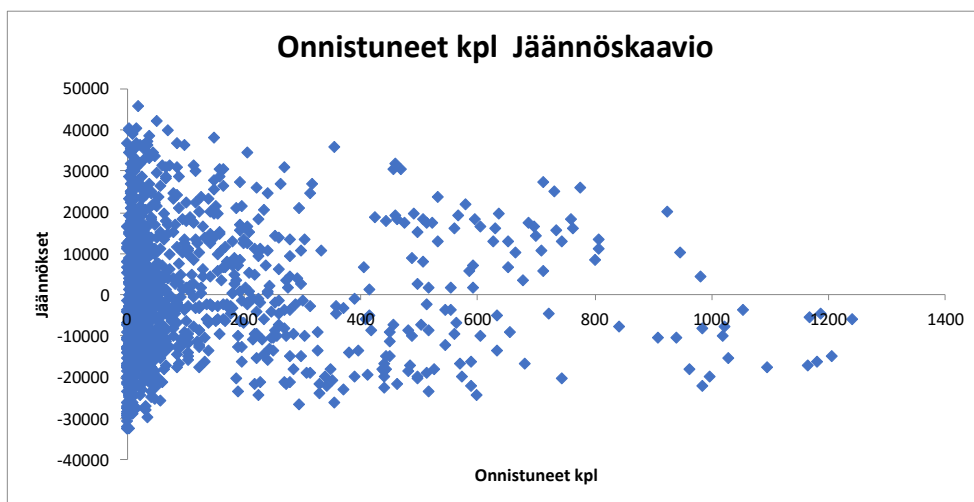
Mallin täytyy lisäksi täyttää tiettyjä edeltävyyssehtoja ollakseen tilastollisesti merkitsevää. Lineaaristen regressiomallien edeltävyys ehdot edellyttävät, että selittävän ja selitettävän muuttujan välillä on lineaarinen riippuvuus, jäännösten varianssi on yhtä suuri kaikilla selittävien muuttujien arvoilla, jäännökset noudattavat normaalijakaumaa ja jäännökset ovat toisistaan riippumattomia. Lineaarisuutta ja jäännösten varianssien yhtäsuuruutta voidaan tarkastella jäännöskuvion avulla. (Taanila, 2020, s.11).

Jäännöskuvion muodosta voidaan tulkita edellä mainittuja seikkoja, kuten muuttujien lineaarista yhteyttä ja jäännösten varianssia. Siistissä jäännöskuviossa pisteet ovat jakaantuneet ilman säännönmukaisuutta. Mikäli riippuvuus ei ole lineaarinen, niin se yleensä näkyy pisteiden säännönmukaisuutena. (Taanila, 2020, s.13).

Kuvio 13:n neljännessä ruudussa on vastaava tilanne, kuin työni aineistossa ennen poikkeavien havaintojen poistamista. Kuvion 14 perusteella voidaan todeta, että työni muuttujien välillä on lineaarinen riippuvuus ja että jäännösten varianssi on pääasiassa yhtä suuri kaikilla selittävän muuttujan arvoilla. Kuvio kapenee hieman korkeilla Snowfox.AI:n onnistumismäärillä, joka viittaisi siihen, että malli olisi hieman tarkempi suurilla arvoilla. Toisaalta havaintojen määrä korkeilla onnistumismäärillä on myös huomattavasti pienempi, joten en vetäisi tästä liian suorja johtopäätöksiä.



Kuvio 13. Esimerkkejä jäännöskuvioiden tulkinnasta (Taanila, 2020, s.13).



Kuvio 14. Opinnäytetyön aineiston regressioanalyysin jäännöskaavio.

#### 4.6 Validiteetti ja reliabiliteetti

Mittausten onnistumista arvioidaan yleensä validiteetin ja reliabiliteetin avulla. Validiteetti ilmaisee missä määrin on kyetty mittaamaan juuri sitä, mitä pitikin mitata. (Nummenmaa et al., 2019, s. 20). Validiteetti ilmaisee, kuinka hyvin tutkimuksessa käytetty mittausmenetelmä mittaa juuri sitä tutkittavan ilmiön ominaisuutta, jota on tarkoituskin mitata (Tilastokeskus, 2024b). Reliabiliteetilla tarkoitetaan mittarin luotettavuutta, eli mittausvirheettömyyttä. Tutkimuksen reliabiliteetti on suuri, mikäli eri mittauskerroilla saadaan samanlaisia tuloksia samasta aineistosta. (Nummenmaa et al., 2019, s.20). Reliabiliteettia voidaan arvioida esimerkiksi toistomittauksilla (Tilastokeskus, 2024c).

Tässä opinnäytetyössä tutkimusmenetelmäksi valikoitui lineaarinen regressioanalyysi ja mitattaviksi muuttujiksi ostolaskujen käsittelyn tuottavuus ja tekoälyn oikein tiliöimien laskujen määrä. Lineaarista regressiomallia voidaan pitää validina mittausmenetelmänä vastaamaan opinnäytetyöni tutkimuskysymykseen ”miten Snowfox.AI-ostolaskutekoäly vaikuttaa Palkeiden ostolaskujen käsittelyn tuottavuuteen”. Regressioanalyysin avulla saatiin uskottavia tuloksia, kuinka suurta osaa tekoäly selittää tuottavuudesta ja kuinka paljon tekoälyn suoritukset lisäävät tuottavuutta, joten menetelmän voidaan arvioida olleen validi tutkimusmenetelmä opinnäytetyössäni.

Opinnäytetyöni reliabiliteettia voidaan arvioida esimerkiksi toistomittausten kautta. Aineisto on suoraan johdettu lähdejärjestelmistä ja aineisto on varsinkin Snowfox.AI:n aineiston osalta luonteeltaan ns. kylmää dataa, joten toistomittauksilla ei olisi muuta mahdollisuutta, kuin saada samoja lopputuloksia. Tältä osin tuloksia voidaan pitää mittausvirheettöminä. Tuottavuusaineiston poikkeavat havainnot viittasivat osittain virheellisiin työajan kirjauksiin. Poikkeavien havaintojen sisällyttäminen regressioanalyysiin vaikutti mallin selityskertoimeen noin yhden prosenttiyksikön verran verrattuna tuloksiin, joiden aineistosta poikkeavat havainnot oli poistettu. Työajan kirjaamiseen ei Palkeilla ole absoluuttisen tarkkaa tapaa, vaan kirjaukset eri toiminnoille tekee työntekijä itse 15 minuutin tarkkuudella, joten hajontaa absoluuttisessa kirjaustarkkuudessa on aina. Kuitenkin poikkeavien havaintojen sisällyttäminen toi

ainoastaan noin prosenttiyksikön vaikutuksen lopputuloksiin. Lisäksi samaa työajan kirjaamisen tapaa Palkeilla käytetään kaikkeen muuhunkin tuottavuuslaskentaan, eikä tätä opinnäytetyötä varten ole käytetty jotain eri menetelmää, joten tämän työn tuloksia voidaan pitää samassa kontekstissa olevina ja siten luotettavina.

Huomionarvoista kuitenkin on, että työajan kirjaamisen tavat vaikuttavat tuottavuuslaskentaan. On eri tutkimuksen tehtävä miettiä sitä, mitä kirjauksia tuottavuuslaskentaan sisällytetään, miten eri toiminnoille kirjattava työaika ohjeistetaan tai löytyisikö järjestelmistä mahdollisuutta seurannan automaatioon. On myös arvioitava saisiko tarkemmalla tai absoluuttisemmalla työajan seurannalla merkittävästi tarkkuutta tuottavuuslaskentaan ja olisiko sillä merkitystä käytännön tekemiseen tai prosessien kehittämiseen. Oma arvioni on, että nykyinen tapa on riittävän tarkka käytännön kannalta, muttei tarkemmasta mitaamisesta olisi myöskään mitään haittaa.

## 5 JOHTOPÄÄTÖKSET

Regressioanalyysin tulokset osoittavat Snowfox.AI-ostolaskutekoälyn selittävän tuottavuutta tietyiltä osin. Asia on toki pääteltävissä jo lähtökohtaisesti, mutta vaikuttavuuden kokoluokka, tai prosentuaalinen osuus ei ole tarkasti pääteltävissä ilman tarkempaa analyysiä. Lineaarisen regressioanalyysin mukaan Snowfox.AI selittää tuottavuuden vaihtelusta noin 8 % ja yksi Snowfox.AI:n oikein ennustama tiliöinti parantaa tuottavuutta noin 23 ostolaskua / tehollinen HTV. Lineaarisen regressioanalyysin avulla olisi mahdollista tarkastella useamman selittävän muuttujan vaikutusta tuottavuuteen, jolloin saataisiin enemmän tietoa tuottavuuteen vaikuttavista tekijöistä ja kenties hyödynnettyä tuloksia vielä paremmin. Useamman selittävän muuttujan tuominen AMK-opinnäytetyön laajuiseen työhön osoittautui kuitenkin haastavaksi, koska se olisi kasvattanut työn teoria- sekä empiriaosuutta vaadittua suuremmaksi.

Opinnäytetyöni kaltaisen regressioanalyysin tulokset olisivat vielä paremmin hyödynnettävissä, mikäli vastaavia tuloksia saataisiin muodostettua jatkuvasti ja automaattisesti. Tutkimusaineiston yhdistely ja koostaminen käsin Excelissä tutkittavaan muotoon on varsin tehotonta työtä jatkuvaan seurantaan. Useamman selittävän muuttujan mallilla saataisiin yksityiskohtaista seurantaa muidenkin tekijöiden, kuin AI-tiliöinnin vaikutuksista. Tämä mahdollistaisi vaikuttamisen niihin tekijöihin, joissa tuottavuuspotentiaalia olisi vielä hyödyntämättä.

Yksittäisen tarkastelujakson tuloksena työni antaa tilannekuvan AI-tiliöinnin käyttöönotosta tarkastelujakson loppuun saakka, joten vertailukohtaa siihen, mistä tultiin ja mihin ollaan menossa ei tällä tutkimuksella selvitetty. Toisaalta työni on kartoittava tutkimus, jonka tarkoituksena on luoda uusia näkökulmia tarkasteltavaan asiaan ja auttaa suunnittelemaan tarkemmin kohdennettuja tutkimuksia tulevaisuudessa. (Abbadia, 2023). AI-tiliöinnin tuottavuusvaikutusten seurantaan liittyvän sovelluksen kehittäminen, useamman selittävän muuttujan malli tai esimerkiksi tuottavuuteen vaikuttavien tekijöiden laajempi analysointi kvantitatiivisella tai kvalitatiivisella menetelmällä voisivat olla potentiaalisia aiheeseen liittyviä aiheita esimerkiksi tuleville opinnäytetöille.

Esittäessäni opinnäytetyöni tuloksia toimeksiantajalle syntyi keskustelua tulosten hyödynnettävyydestä. Tekoälyn vaikutusta ostolaskujen käsittelyn kokonaistuottavuuteen pidettiin kiinnostavana näkökulmana. Asiakaskohtaisesti eriteltyjen tulosten avulla voitaisiin erottaa sellaiset asiakkaat, joilla tekoälyn osuus tuottavuuteen on onnistuttu saamaan korkeaksi. Korkean tekoälytuottavuuden saavuttaneiden asiakkaiden tulosten avulla voitaisiin analysoida, onko kyseisen asiakkaan ostolaskujen käsittelyn tavoissa jotain tekijöitä, joita voisi hyödyntää myös enemmän manuaalista työtä vaativien asiakkaiden kanssa. Snowfox.AI:n onnistumisprosentti toimii laskutusperusteena palvelulle, joten olisi kaikkien etu, mikäli tekoälyn tuottavuutta saataisiin parannettua.

Keskustelua heräsi myös siitä, kuinka ostolaskujen manuaaliselle käsittelylle kohdennetut työajat olivat vähentyneet huomattavasti tekoälyn käyttöönoton jälkeen. Vaikkei tiliöintiennuste ole sataprosenttisesti oikein, nopeuttaa Snowfox.AI:n laskulle luomat tiliöintirivit käsittelyä huomattavasti, vaikka tiliöinti

vaatisi korjaamista. Keskustelua herätti myös nykyiset työajan kirjaamisen tavat, jotka eivät välttämättä ole optimaalisia tuottavuuslaskennan kannalta. Työajan kirjaamisen tavat koettiin osittain vanhentuneiksi ja työajan kirjaamisen osaamisen tasoa palveluryhmissä pohdittiin. Virheelliset työajankirjaukset tai kirjaamatta jättämiset selittivät osan tämän opinnäytetyön aineiston poikkeavista havainnoista.

Keskustelussa nousi myös esiin EU:n tuleva tekoäylaki ja sääntelyn vaikutukset tekoälyn käyttöön taloushallinnon prosesseissa. EU:n sääntelyehdotuksen tavoitteena on tarjota tekoälyn kehittäjille ja käyttäjille selkeät vaatimukset ja velvoitteet, jotka koskevat tiettyjä tekoälyn käyttötarkoituksia (Euroopan komissio, 2024). Tekoälyteknologia on nopeasti kehittyvä ala, joten EU:n sääntelyehdotus ei vielä ota konkreettisesti kantaa asioihin, jotka vaikuttaisivat Palkeiden toimintaan, mutta ehdotuksen kehittymistä seurataan mielenkiinnolla.

## 6 POHDINTA

Opinnäytetyön tavoitteena oli kartoittaa tilastollisella menetelmällä keväällä 2022 käyttöön otetun Snowfox.AI-ostolaskutekoälyn vaikutusta Palkeiden ostolaskujen käsittelyn tuottavuuteen. Lineaarisen regressioanalyysin avulla tämä tavoite sekä uskottavat tulokset saavutettiin. Tulosten myötä saatiin tieto ostolaskutekoälyn vaikutuksista tuottavuuteen palvelun käyttöönoton ja tutkimusajankohdan väliseltä tarkastelujaksolta. Kartoittavan tutkimuksen avulla saatiin tietoa tulosten hyödynnettävyyteen esimerkiksi tulevien opinnäytetöiden suunnittelussa.

Opinnäytetyö oli itselleni ensimmäinen käytännön sovellus lineaarisen regressiomallin käyttöön, joten työ oli kokonaisuudessaan myös uudenlainen oppimiskokemus. Opin itselleni uuden tilastollisen mallin hyödyntämisen lisäksi hahmottamaan tuottavuuden käsitettä organisaatiossa paremmin sekä sen, kuinka moni asia tuottavuuteen lopulta vaikuttaa. Alun perin tarkoitukseni oli

sisällyttää muitakin selittäviä muuttujia malliin, mutta opinnäytetyön laajuuden, opinnoista valmistumisen aikataulun sekä itselleni uuden tilastollisen menetelmän opetteluun vaaditun ajan huomioon ottaen päädyin tarkastelemaan ai-  
hetta ainoastaan koneoppivan tekoälyn näkökulmasta. Aihe on ajankohtainen ja tulee vaikuttamaan taloushallinnon prosesseihin tulevina vuosina merkittävästi, joka teki aiheen tutkimisesta myös mielenkiintoista.

Vaikka AI-tiliöinnistä on huomattava käytännön apu ostolaskujen käsittelyssä, opinnäytetyöni tulosten perusteella se ei kuitenkaan vielä selitä kovin merkittävää osaa ostolaskujen käsittelyn kokonaistuottavuudesta. Ostolaskujen käsittely vaatii edelleen huomattavan määrän käsityötä, eikä tekoälyratkaisut tai automatiikka vielä pysty syrjäyttämään ihmistä ostolaskujen käsittelyn prosesseissa. Koneoppiva tekoäly oppii parhaiten, mikäli toimittajan laskut ovat toistuvasti samankaltaisia. Silloin tekoäly oppii helpoiten laskujen sisällöstä sekä niiden tiliöinneistä ja ennusteiden onnistumisprosentti paranee. Tosielämässä laskut ovat kuitenkin keskenään hyvin erilaisia ja toimittajilta tulee laskuja erilaisista palveluista ja tuotteista eri asiakkaille ja näitä laskuja kirjataan eri projekteille ja seurantakohteille. Laskujen erittelyt voivat myös olla erillisiä varsinaisista laskudokumenteista, joka myös haastaa tekoälyn oppimista, koska tekoäly ei pysty tunnistamaan mitä on laskutettu. Potentiaalia uudessa teknologiassa kuitenkin on ja uusi teknologia tulee muuttamaan taloushallinnon alaa tulevina vuosina.

Opinnäytetyö noudattaa ammattikorkeakoulujen opinnäytetöille määritellyjä eettisiä suosituksia. Eettiset suositukset on otettu huomioon kaikissa opinnäytetyön vaiheissa. Opinnäytetyössä esitetyt tiedot on esitetty avoimesti ja ilman vääristelyä. Opinnäytetyö ei sisällä arkaluonteista aineistoa tai tietoa ja tutkitavat asiakkaat ovat tunnistamattomissa, kuten opinnäytetyön toimeksiantajan kanssa on sovittu. Lähteet on merkitty asianmukaisesti ja lainaukset on merkitty tekstiin selkeästi. Opinnäytetyön tarkoituksena on ollut tuottaa luotettava tieteellinen tutkimus.

## LÄHTEET

- Abbadia, J. (25.10.2023) Tutkivan tutkimuskysymyksen esimerkit. Haettu 12.11.2023 osoitteesta <https://mindthegraph.com/blog/fi/tutkimus-tutkimuskysymys-esimerkkeja/>
- Bonnier Pro – Parhaat käytännöt liiketoimintaasi. (2024). Talous ja rahoitus – Järjestelmät ja prosessit. Haettu 17.01.2024 osoitteesta <https://bonnierpro-fi.lillukka.samk.fi/fi/>
- Briox. (2024). Tiliöinti. Haettu 28.01.2024 osoitteesta <https://briox.fi/sana-kirja/tiliointi/>
- Euroopan komissio. (2024). Tekoälylaki. Haettu 18.02.2024 osoitteesta <https://digital-strategy.ec.europa.eu/fi/policies/regulatory-framework-ai>
- Heikkilä, T. (2014). Tilastollinen tutkimus. Edita Publishing Oy.
- Heikkinen, N. (15.9.2023). Palkeiden taloushallinnon asiantuntija Niko Heikkisen sähköposti.
- Herrmann, C. & Tarren, D. (2012). Tuottavuus julkisella sektorilla. European public service union. Haettu 11.11.2023 osoitteesta [https://www.epsu.org/sites/default/files/article/files/Productivity\\_Final\\_Draft\\_FI\\_rev.pdf](https://www.epsu.org/sites/default/files/article/files/Productivity_Final_Draft_FI_rev.pdf)
- Iltsanommat. (29.04.1998). Taloushallinto kokee vallankumouksen – Paperiton kirjanpito on tulevaisuuden tekemistä, ei historian tallentamista. Haettu 16.01.2024 osoitteesta <https://www.is.fi/taloussanommat/art-2000001285038.html>
- Kaarlejärvi, S. & Salminen, T. (2018). Älykäs taloushallinto – Automaation aika. Alma Talent Oy.
- Kananen, H. & Puolitaival, H. (2019). Tekoäly – Bisneksen uudet työkalut
- Lahti, S. & Salminen, T. (2014). Digitaalinen taloushallinto. Alma Talent Oy.
- Laki Valtion talous- ja henkilöstöhallinnon palvelukeskuksesta 179/2019 muutoksineen. <https://www.finlex.fi/fi/laki/ajantasa/2019/20190179>
- Listenmaa, J. (2023). Laita tieto töihin – Tiedolla johtamisen käsikirja. Alma Talent Oy.
- Microsoft. (2023). Ruutu- ja janakaavion luominen. Haettu 15.01.2023 osoitteesta <https://support.microsoft.com/fi-fi/office/ruutu-ja-janakaavion-luominen-62f4219f-db4b-4754-aca8-4743f6190f0d>
- Neilimo, K. & Uusi-Rauva, E. (2017). Johdon laskentatoimi (6.–13. painos). Edita Publishing Oy.



Nummenmaa, L., Holopainen, M. & Pulkkinen, P. (2019). Tilastollisten menetelmien perusteet. Sanoma Pro Oy.

Palkeet. (2023a). Uudistuvan julkishallinnon kumppani talous- ja henkilöstöpalveluissa. Haettu 16.9.2023 osoitteesta <https://www.palkeet.fi/palkeet.html>

Palkeet. (2023b). Palkeet lyhyesti [intranet].

Palkeet. (2023c). Yhdessä kehittäen. Haettu 08.11.2023 osoitteesta <https://www.palkeet.fi/yhdessa-kehittaen.html>

Palkeet. (2024a). Menotositteiden tiliöinti [intranet].

Palkeet. (2024b). Ostolaskut ja automaatioaste. Microsoft PowerBI-raportti [intranet]

Palkeet. (2024c). Talousarviokirjanpidon perusteet. [PowerPoint-diat, intranet].

Palkeet. (2024d). Tilinpäätös vuodelta 2022. Haettu 05.02.2024 osoitteesta <https://www.palkeet.fi/media/sivujen-liitetiedostot/tilinpaatos-2022.pdf>

Riihos. (2024). Tiliöinti – Mitä tarkoitetaan tiliöinnillä kirjanpidossa? Haettu 17.01.2024 osoitteesta <https://riihos.fi/tiliointi/>

Snowfox.AI. (2024). Automatisoi ostolaskujen tiliöinti ja kiertoalaitto tekoälyllä. Haettu 16.01.2024 osoitteesta <https://www.snowfox.ai/fi/>

Sivonen, M. (2017). Dokumenttianalyysi tutkimusmentelmänä terveystieteissä. [Kuvaileva kirjallisuuskatsaus, Itä-Suomen yliopisto]. Haettu 11.11.2023 osoitteesta [https://erepo.uef.fi/bitstream/handle/123456789/17863/urn\\_nbn\\_fi\\_uef-20170525.pdf](https://erepo.uef.fi/bitstream/handle/123456789/17863/urn_nbn_fi_uef-20170525.pdf)

Taanila, A. (28.9.2020). Lineaariset regressiomallit. Haettu 14.01.2024 osoitteesta <http://taanila.fi/regressio.pdf>

Tarkka, H. (2010). Julkisen hallinnon tuottavuus ja tietojärjestelmien yhteentoimivuus. Taloustieteellinen yhdistys. Haettu 29.10.2023 osoitteesta <https://www.taloustieteellinenyhdistys.fi/kak12010/>

Tietoarkisto. (2024a). Kovarianssi ja korrelaatio. Haettu 18.01.2024 osoitteesta <https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvanti/korrelaatio/korrelaatio/#kovarianssin-standardointi-eli-korrelaatiokerroin>

Tietoarkisto. (2024b). Regressioanalyysi. Haettu 18.01.2024 osoitteesta <https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvanti/regressio/analyysi/>

Tietoarkisto. (2024c). Regressiomallin arviointi. Haettu 19.01.2024 osoitteesta <https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvanti/regressio/arviointi/>

Tietoarkisto. (2024d). Hypoteesien testaus. Haettu 24.02.2024 osoitteesta <https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvanti/hypoteesi/testaus/>

Tilastokeskus. (2024a). Henkilötyövuosi. Haettu 17.01.2024 osoitteesta <https://www.stat.fi/meta/kas/henkilotyovuosi.html>

Tilastokeskus. (2024b). Validiteetti. Haettu 24.02.2024 osoitteesta <https://www.stat.fi/meta/kas/validiteetti.html>

Tilastokeskus. (2024c). Reliabiliteetti. Haettu 24.02.2024 osoitteesta <https://www.stat.fi/meta/kas/reliabiliteetti.html>

Valli, R. (2015). Johdatus tilastolliseen tutkimukseen. PS-kustannus.

Valtiokonttori. (2024). Kirjanpidon tilit ja tilikartta. Haettu 28.01.2024 osoitteesta <https://www.valtiokonttori.fi/maaraykset-ja-ohjeet/kirjanpidon-tilit-ja-tilikartta-11/>

Valtiovarainministeriö. (2023a). Tilinpäätöskannanotto – Valtiovarainministeriön kannanotto vuoden 2022 tilinpäätöksestä. Haettu 29.10.2023 osoitteesta [https://www.palkeet.fi/media/suunnittelu\\_ja\\_seuranta\\_netra/vm-tilinpaatos-kannanotto-2022.pdf](https://www.palkeet.fi/media/suunnittelu_ja_seuranta_netra/vm-tilinpaatos-kannanotto-2022.pdf)

Valtiovarainministeriö. (2023b). Valtion taloushallinto. Haettu 16.9.2023 osoitteesta <https://vm.fi/hallintopolitiikka/valtion-taloushallinto>