DATAMALLI PÄÄTÖKSENTEON TUKENA: CASE ESPOON KAUPUNGIN KOULUKULJETUKSET

Lappeenrannan–Lahden teknillinen yliopisto LUT

Tuotantotalouden diplomityö

2024

Juha Forsström

Tarkastajat: Dosentti Kalle Elfvengren

Dosentti Lasse Metso

TIIVISTELMÄ

Lappeenrannan–Lahden teknillinen yliopisto LUT

LUT Teknis-luonnontieteellinen

Tuotantotalous

Juha Forsström

Datamalli päätöksenteon tukena: Case Espoon kaupungin koulukuljetukset

Tuotantotalouden diplomityö

2024

105 sivua, 30 kuvaa, 12 taulukkoa ja 1 liite

Tarkastajat: dosentti Kalle Elfvengren ja dosentti Lasse Metso

Avainsanat: Data päätöksenteon tukena, tiedolla johtaminen, datan visualisointi, julkinen sektori, koulukuljetukset

Tietojohtaminen on ilmiönä ajankohtainen niin yksityisellä kuin julkisella sektorillakin. Dataa on yhä enemmissä määrin saatavissa ja sitä pyritään hyödyntämään siten, että siitä saadaan hyödyllistä tietoa päätöksenteon tueksi. Organisaatioiden johtamista toteutetaan edistämällä kykyä luoda arvoa tiedolla ja osaamisella. Pyrkimyksenä on olemassa olevan ja sinne saatavissa olevan datan hyödyntäminen sekä myös sen saatavuuden varmistaminen.

Tämän opinnäytetyön toimeksiantajana toimii Espoon kaupungin Logistiikka -vastuuyksikkö. Opinnäytetyön tavoitteena on selvittää miten muut kunnat hyödyntävät dataa koulukuljetuksia koskevan päätöksenteon tukena. Selvityksen lisäksi tavoitteena on rakentaa Power BI -datamalli kohdeorganisaation käyttöön työkaluksi, mitä vastuuyksikkö voi hyödyntää koulukuljetuksia koskevassa päätöksenteossa. Tarve työhön tuli toimeksiantajalta, mikä ei ole aikaisemmin teettänyt samanlaista selvitystä. Datan hyödyntämisen kasvattaminen on ajankohtainen osana kohdeorganisaation strategiaa ja siinä korostuvaa tiedolla johtamista.

Tutkimus on toteutettu kvalitatiivisin menetelmin konstruktiivisena tapaustutkimuksena. Konstruktion rakentamiseen on hyödynnetty kirjallisuuskatsauksessa käsiteltyä teoriaa. Työn empiirinen aineisto koostuu kohdeorganisaation koulukuljetusdatasta sekä haastattelujen tuloksista. Datamallin rakentamisessa on sovellettu OSEMN-prosessimallia.

Työn tuloksena on tuotettu ehdotus Power BI -datamallista, mikä on rakennettu monivaiheisen prosessin tuotoksena. Syntynyt konstruktio on otettu käyttöön kohdeorganisaatiossa, joka jatkossa voi kehittää ja laajentaa tuotosta muuttuvien tietotarpeidensa sekä liiketoimintakentän mukaan.

ABSTRACT

Lappeenranta–Lahti University of Technology LUT

School of Engineering Science

Industrial Engineering and Management

Juha Forsström

Data model to support decision making: Case of school transportation in the city of Espoo

Master's thesis

2024

105 pages, 30 figures, 12 tables and 1 appendice

Examiners: Docent Kalle Elfvengren and Docent Lasse Metso

Keywords: Data in decision making, information management, data visualization, public sector, school transportation

Knowledge management is a current phenomenon in both the private and public sectors. Increasingly, more data is available, and there is an effort to utilize it to extract useful information to support decision-making. Organizations are managed by promoting the ability to create value through knowledge and expertise. The aim is to leverage the existing data and ensure its availability.

The commissioner of this thesis is the city of Espoo's logistics unit that is responsible for school transportation in the city. The objective of the thesis is to investigate how other cities utilize data to support decision-making regarding school transportation. In addition to the investigation, the goal is to build a Power BI data model for the target organization to use as a tool that the logistics unit can utilize in decision-making about school transportation. The need for this work arose from the commissioner, which has not previously conducted a similar investigation. Data utilization plays increasingly big part of the target organization's strategy, which emphasizes knowledge-based management.

The research has been conducted using qualitative methods as a constructive case study. The construction has utilized the theory discussed in the literature review. The empirical data of the work consists of school transportation data from the target organization and the results of interviews. The OSEMN process model has been applied in building the data model.

The result of the work is a proposal for a Power BI data model comprising six report pages. The output has been created through a multi-phase process. The resulting construction has been implemented in the target organization, which can further develop and expand the output according to its changing information needs and business environment.

ALKUSANAT

Matkani Espoon kaupungin Logistiikka -yksikössä on ollut opettavaista. Tunnen, että olen päässyt etuoikeutetusti en vain kokemaan miten henkilökuljetukset hoituvat julkisella sektorilla Suomen toiseksi suurimmassa kaupungissa vaan myös näkemään organisaation merkittäviä rakenteellisia muutoksia. Kunnalliset uudistukset, sopimusmuutokset, talousjärjestelmän sekä toiminnanohjausjärjestelmän vaihtumiset ovat isoja muutoksia, mitkä ovat tapahtuneet aikaisemmin ja mitkä tapahtuvat myös tulevaisuudessa muuttaen merkittävästi toiminnan luonnetta organisaatiossa ja mihin sen on sopeuduttava. Tiedolla johtaminen on Espoon kaupungin strategian keskiössä, joten matkani suunta on toiminnan ohjaaminen entistä datalähtöisemmäksi. Tämä työ on toteutettu Logistiikka -yksikön toiminnan datalähtöisempään suuntaan menemisen tueksi. Haluan kiittää työnantajaani ja ohjaajia mahdollisuudesta tehdä diplomityö mielenkiintoisesta ja ajankohtaisesta aiheesta.

30.6.2024 Helsinki

Lyhenteet

BI	liiketoimintatiedon hyödyntäminen tai hallinta (Business Intelligence)
DAX	Data Analysis Expressions
DIKW	Data-Information-Knowledge-Wisdom -hierarkia
ETL	tarkoittaa datan siirtämisen (extract), muokkaamisen (transform) ja lataamisen (load) prosessia
KOTO	Kasvun ja oppimisen toimiala
OSEMN	viitekehys datatieteen prosessille. Akronyymi sanoista Obtain, Scrub, Explore, Model ja iNtrepret

Sisällysluettelo

Tiivistelmä

Abstract

Alkusanat

Symboli- ja lyhenneluettelo

1 Johdanto ... 3
1.1 Työn tausta ja motivaatio ... 3
1.2 Tavoitteet, tutkimuskysymykset ja rajaus ... 4
1.3 Tutkimuksen toteutus ja raportin rakenne ... 5
2 Datan hyödyntämisestä ja mallintamisesta ... 7
2.1 Datasta tiedoksi ja tiedolla johtaminen ... 7
2.1.1 Datasta hyödynnettäväksi tiedoksi ... 7
2.1.2 Tiedolla johtaminen ... 9
2.2 Datalähtöisyys päätöksenteon tukena ... 11
2.2.1 Data-analytiikka ... 11
2.2.2 Business Intelligence - Liiketoimintatiedon hyödyntäminen ... 13
2.2.3 Perinteinen ja itsepalvelu BI ... 15
2.2.4 Datan visualisointi: tieto näkyväksi ... 16
2.3 Datamallin kehittämisen keskeinen terminologia ... 19
2.3.1 Datan mallintaminen ... 19
2.3.2 Datamallin laajentaminen DAX:illa ... 24
3 Tutkimus – Espoon kaupungin koulukuljetukset ... 26
3.1 Tutkimusmetodologia ja menetelmät ... 26
3.2 Kohdeorganisaatio ... 30
3.2.1 Logistiikka (vastuuyksikkö) ja henkilökuljetukset ... 30
3.2.2 Koulukuljetukset Espoossa ... 31
3.2.3 Järjestelmätoimittajan kuvaus ... 32
3.2.4 Tietojohtaminen, Excel- ja Power BI-raportointi Espoossa ... 34
3.3 Osa I: Tutkimuksen haastattelut ... 37

3.3.1 Haastattelujen toteutus 37
3.3.2 Haastattelujen tulokset 40
3.4 Osa II: Datamalli 57
3.4.1 Sovelluksena Power BI 58
3.4.2 Tavoitteen määrittäminen 59
3.4.3 Data-aineiston hankkiminen ja kuvaus 61
3.4.4 Datan puhdistus ja muuntaminen 65
3.4.5 Datan tutkiminen 68
3.4.6 Mallinnus ja Datamallin rakentaminen 69
3.4.7 Tulosten tulkinta 94
4 Johtopäätökset 97
5 Yhteenveto 100
Lähteet 101

Liitteet

Liite 1. Haastattelun kysymysrunko

1 Johdanto

Tässä luvussa esitellään diplomityön lähtökohdat ja annetaan yleiskatsaus tutkimuksen aiheesta. Aluksi kartoitetaan työn taustaa ja motivaatiota, minkä jälkeen käsitellään työn tavoitteet sekä esitellään tutkimuskysymykset ja aiheen rajaukset. Lopuksi esitellään tutkimuksen toteutus sekä raportin rakenne.

1.1 Työn tausta ja motivaatio

Tietojohtaminen on ilmiönä ajankohtainen niin yksityisellä kuin julkisella sektorillakin ja siihen liittyvät termit esiintyvät usein julkisessa keskustelussa sekä erilaisissa yhteiskunnan ja yksityisen sektorin strategioissa ja julkilausumissa (Helander et al. 2020). Organisaatioiden johtamista toteutetaan siten, että edistetään kykyä luoda arvoa tiedolla ja osaamisella. Pyrkimyksenä on organisaatiossa olemassa olevan ja sinne saatavissa olevan tiedon hyödyntäminen sekä myös saatavuuden varmistaminen. Dataa käytetään, jotta voidaan poimia oivalluksia, tunnistaa trendejä ja tehdä päätöksiä eri tasoilla. Analytiikan avulla organisaatiot voivat tehdä johtopäätöksiä ja ennusteita sekä saada toimeenpantavia oivalluksia asiayhteyteen ja aikarajaan perustavalla tavalla. Päätöksenteko on jatkuva liiketoimintaprosessi, joten analytiikka nopeuttaa sitä ja antaa organisaatioille vauhtia pysyä ajan tasalla muuttuvissa liiketoimintaympäristössä. (Sharma et al. 2023)

Tämän työn tutkimuksen aiheena on datan hyödyntäminen kaupungin koulukuljetusten järjestämisen ja siihen liittyvän päätöksenteon tukena. Kohdeorganisaatio tutkimukselle on Espoon kaupungin Logistiikka -vastuuyksikkö (jatkossa myös Logistiikka). Logistiikka järjestää kaupungin henkilökuljetuksia kuntalaisille. Näihin henkilökuljetuksiin kuuluvat vuoden 2023 alusta alkaen koulukuljetukset sekä varhaiskasvatuksen kuljetukset. Sote-kuljetukset, kuten vanhusten ja vammaisten kuljetukset, kuuluivat myös aikaisemmin kaupungin hallinnoitavaksi, mutta ne siirtyivät Sote-uudistuksen myötä Länsi-Uudenmaan hyvinvointialueen järjestämisvastuuseen vuoden vaihteessa.

Ennen vuoden 2023 alussa voimaan astunutta Sote-uudistusta Espoon henkilökuljetuksien liittyvistä luvuista toteutettiin kuukausiseurantaa ja raportointia vanhan kuljetustentilausjärjestelmän datalla. Vuonna 2022 logistiikan kuljetustenhallintajärjestelmän vaihtui ja nyt uuden järjestelmän myötä logistiikkatiimin on tarkoitus parantaa henkilökuljetuksien lukujen seurantaa tukemaan päätöksentekoa niin Ad hoc -tehtävissä kuin myös laajemmissa tehtäväkokonaisuuksissa kuten kilpailutuksissa. Toistaiseksi henkilökuljetuksista pidetään kuukausittaista seurantaraportointia asetettujen päämittareiden avulla, jotka ovat kuljetusten määrä, asiakkaiden määrä, reklamaatioiden suhde kuljetusten määrään.

Espoon kaupungin määrittelemän tarinansa tai lyhyemmin Espoo-tarinan mukaisesti johtamista kehitetään, tiedolla, data-analyysilla ja vaikuttavuudella johtamista. Espoon tavoitteena on parantaa paleluiden vaikuttavuutta, nopeuttaa palveluprosesseja sekä saavuttaa kustannussäästöjä. (Espoon kaupunki 2023g)

1.2 Tavoitteet, tutkimuskysymykset ja rajaus

Tämän tutkimuksen tavoitteena on parantaa Espoon kaupungin Logistiikkayksikön hallinnoimien koulukuljetusten järjestämistä sekä siihen liittyvää päätöksentekoa. Tavoitteeseen pääsemisenin toteutetaan haastattelemalla muiden kaupunkien asiantuntijoita ja havaitsemalla niistä hyviä dataan liittyviä käytänteitä sekä rakentamalla datamalli haastatteluiden tuloksista, verkkoaineistosta löytyvien hyödyllisiksi havaittujen aineistojen perusteella sekä yksikön tarpeiden mukaisesti. Datamallin on tarkoitus helpottaa koulukuljetuksiin liittyvää päätöksentekoa.

Tutkimus rajataan henkilökuljetuksiin ja yhä tarkemmin vielä koulukuljetuksiin, mitkä ovat opetuksen toimialan myöntämällä kuljetuspäätöksellä järjestetty koulutaksein oppilaan koulun sekä kodin tai hoitopaikan välillä. Rajaus on tehty koskemaan kyseisiä kuljetuksia, koska ne ovat nykyisellään merkittävin osa Espoon kaupungin henkilökuljetusten kokonaisuutta niin kustannuksellisesti sekä asiakas- ja kuljetusmäärällisesti. Täten tilausbussikuljetukset, esimerkiksi oppilaiden kuljetukset koulujen väistötiloihin tai kouluretkille, sekä matkakorttipäätöksellä julkisilla kulkuneuvoilla kulkevat oppilaat on rajattu pois tutkimuksesta. Rajauksesta huolimatta, on tarkoitus, että työn tuloksia voidaan hyödyntää tarvittaessa myös

muiden tyyppisten henkilökuljetusten tarkastelun apuna, muun muassa datamallissa suodatusta keventämällä. Lisäksi datamalli rajataan vuoden 2023 tammikuun ja sen jälkeiseen dataan, koska uusi kuljetustenhallintajärjestelmä otettiin silloin täysvaltaisesti käyttöön ja vanha järjestelmä poistui käytöstä.

Työn tutkimuskysymykset ovat:

1. Miten Espoon kaupunki voi tukea päätöksentekoa logistiikan koulukuljetusten osalta paremmin dataa hyödyntämällä?
2. Miten Suomen muut kunnat/kaupungit hyödyntävät dataa koulukuljetusten hallinnoinnissa ja mitkä ovat hyviä käytäntöjä, joita Espoon kaupunki voisi ottaa käyttöön toiminnassaan?

1.3 Tutkimuksen toteutus ja raportin rakenne

Tämä tutkimus toteutetaan konstruktiivisena tapaustutkimuksena, joka jakautuu teoreettiseen kirjallisuus osuuteen sekä empiiriseen tapaus osioon. Päälukuja on yhteensä viisi, jotka ovat Johdanto, Datan hyödyntämisestä ja mallintamisesta, Tutkimus – Espoon kaupungin koulukuljetukset, Johtopäätökset sekä Yhteenveto. Taulukossa 1 on esitetty työn rakennekaavio. Ensimmäinen luku esittelee työn taustaa ja tarkoitusta kuvaamalla työn tavoitteet, rajaukset ja tutkimuskysymykset. Toinen luku käsittelee tutkimuksen kannalta tärkeitä käsitteitä ja siten taustoittaa työtä. Kolmas luku käsittää työn empiirisen osan ja itse tutkimuksen. Luvussa neljä käsitellään tutkimuksesta saadut johtopäätökset ja esitellään työn tulokset sekä vastaukset tutkimuskysymyksiin. Viimeinen ja kuudes luku käsittää yhteenvedon työstä.

Taulukko 1 Työn rakennekaavio

Input	**Otsikko**	**Output**
Työn tausta ja tarkoitus	1.Johdanto	Työn tavoite, rajaukset ja tutkimuskysymykset
Hakusanat: data, analytics, business intelligence, ETL, DAX, data visualization Tietokannat: LUT Primo, Scopus, Google Scholar	2.Datan hyödyntämisestä ja mallintamisesta	Kirjallisuus datasta, analytiikasta, tiedolla johtamisesta, Business Intelligence:sta, datan mallintamisesta ja siihen liittyvästä teoriasta.
Kohdeorganisaation esittely, Haastattelut ja Datamalli	3.Tutkimus – Espoon kaupungin koulukuljetukset	Tutkimusmetodologian ja menetelmien esittely. Kohdeorganisaation ja sen tietojohtamisen, raportoinnin, koulukuljetuksien sekä niiden vastuuyksikön ja järjestelmätoimittajan kuvaus. Haastattelujen tulokset sekä työn tuotoksena rakennetun datamallin esittely.
Analyysit	4.Johtopäätökset	Tulokset ja vastaukset tutkimuskysymyksiin
Yhteenveto työstä	5.Yhteenveto	Työn yhteenveto

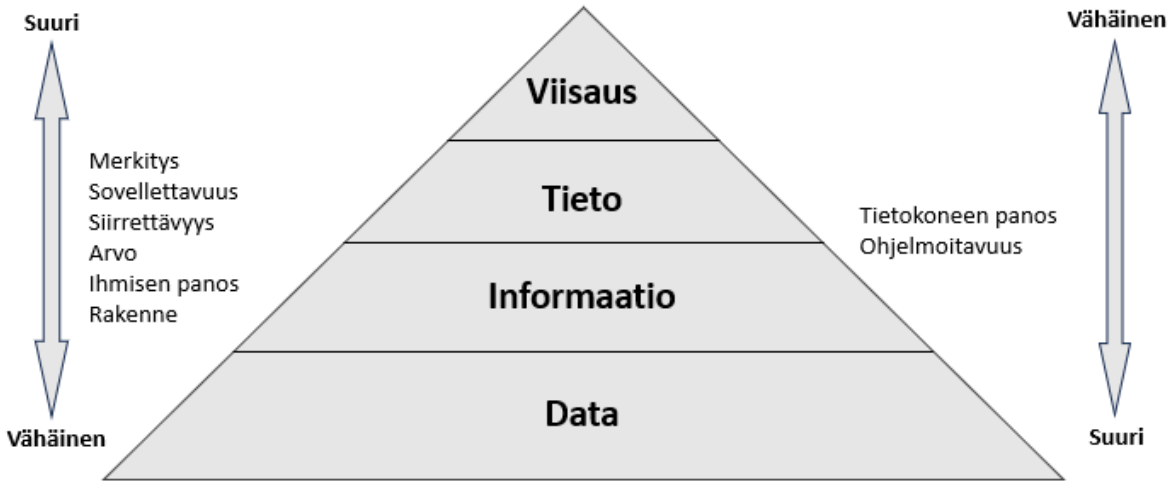
2 Datan hyödyntämisestä ja mallintamisesta

Tässä luvussa tarkastellaan tutkimuksen kannalta keskeiset käsitteet. Aluksi tarkastellaan, miten data muuttuu tiedoksi sekä esitellään työn kannalta olennainen tietojohtamisen käsitteeseen kuuluva tiedolla johtaminen. Lisäksi luku käsittää tarkastelun data-analytiikasta ja sen tasoista sekä business intelligence:stä ja siihen olennaisesti liittyvistä visualisoinneista. Loppuun käsitellään datan mallintamisen keskeinen terminologia. Luvun tavoitteena on antaa lukijalle kattava ymmärrys keskeisistä käsitteistä, jotka ovat olennaisia tutkimuksen toteutuksen ja tulkinnan kannalta.

2.1 Datasta tiedoksi ja tiedolla johtaminen

2.1.1 Datasta hyödynnettäväksi tiedoksi

Vaikka termit data, informaatio ja tieto ovat haastava määritellä yksiselitteisesti, ovat niiden keskinäiset suhteet olennaisia. Liew (2007) toteaa, että monista määrittely-yrityksistä huolimatta ei ole selkeää ja kokonaista kuvaa siitä mitä ne ovat ja mitkä ovat niiden väliset suhteet, kun taas Rowley (2007) korostaa niiden määrittelyn tapahtuvan suhteessa toisiinsa, missä data ja informaatio toimivat tiedon syötteinä ja osana hierarkista rakennetta. Kuvassa 1 esitetty DIKW-malli, joka tunnetaan myös tietopyramidina, kuvaa miten data jalostuu informaatioksi ja tieto edelleen viisaudeksi. Tietopyramidin kaksi alempaa tasoa mahdollistavat tiedon syntymisen mikä taas edelleen voi kehittyä viisaudeksi. Ylemmälle tasolle siirryttäessä merkitys, sovellettavuus, siirrettävyys, arvo, ihmisen panos, sekä rakenne suurenevat. Samalla tietokoneen panos sekä ohjelmoitavuus vähenevät. (Rowley 2007)



Kuva 1 DIKW-malli (mukaillen Rowley 2007)

Data koostuu symboleista, jotka kuvaavat objektien, tapahtumien ja niiden ympäristön ominaisuuksia, ja ovat hyödyllisiä vasta kun ne ovat käyttökelpoisessa muodossa. Informaatio johdetaan datasta vastaa kysymyksiin mitkä alkavat sanoilla kuten *”kuka”*, *”mitä”*, *”milloin”* ja *”kuinka monta”*. Tietojärjestelmät tuottavat, tallentavat, hakevat ja käsittelevät dataa tuottaen informaatio. Tieto puolestaan käsittää osaamisen, joka mahdollistaa informaation muuntamisen ohjeiksi ja voi syntyä joko välittämällä se toiselta henkilöltä, ohjeiden avulla tai kokemuksen kautta. (Rowley 2007) Davenport ja Prusak määrittelevät, että tieto on jäsenneltyjen kokemusten, arvojen, informaation ja oivallusten sekoitus, joka tarjoaa viitekehyksen arvioida uusia kokemuksia ja informaatiota. Tieto syntyy ja sitä sovelletaan tietäjän mielikuvissa. Organisaatiossa tieto on usein sidottu dokumentteihin, rutiineihin, toimintatapoihin ja normeihin (Pentti Salmela 2014). Viisaus on kyky lisätä arvoa tehokkuuden kautta, mikä edellyttää toimijan arvostelukykyä ja henkilökohtaisia eettisiä ja esteettisiä arvoja. (Rowley 2007)

Sherman (2014) kirjoittaa, että datan ja informaation välillä on iso ero, vaikka niitä käytetään usein vaihtokelpoisina termeinä. Data on raakaa, satunnaista ja järjestämätöntä kun taas informaatio on käsiteltyä, jäsenneltyä sekä järjestettyä dataa, mitä käytetään tiedon saamiseksi. Data on lähtötietoja eri käyttöjärjestelmistä ja kun ne syötetään ETL-järjestelmään, niistä tulee informaatiota. Kun tätä informaatiota luetaan raporteista niin siitä tulee tietoa mitä hyödynnetään päätöksissä. (Sherman 2014, s. 8–10) Avain informaation ja tiedon välisen suhteen ymmärtämiseen Liewin (2007) mukaan on tietää missä informaatio sijaitsee,

muistuttaen, että informaatio on pohjimmiltaan viesti mikä on generoitu toiminnoista ja tilanteista. Informaatio sijaitsee tietojärjestelmässä datana ja ihmismielessä tietona. Tämän perusteella datan ja informaation sekä informaation ja tiedon välinen päällekkäisyys on ilmeinen eli toisin sanoen ne ovat samaan aikaan eri tasoilla. Tämä selittää myös miksi käsitteitä usein pidetään keskenään vaihdettavina. Ne eivät ole kuitenkaan vaihdettavissa keskenään niiden hyväksyttyjen erillisten määritelmien osalta.

Tiedon tasojen lisäksi toinen usein käytetty tapa on erottelu hiljaisen tiedon ja eksplisiittisen tiedon välillä. Hiljainen tieto on henkilölle kokemuksen kautta kertynyttä tietämystä, joka on toisinaan tiedostamatonta ja toisinaan tiedostettua. Sitä voidaan kuvata intuitiona ja osaamisena, mikä voi olla vaikea siirtää sanalliseen muotoon. Siten sen siirtäminen toiselle henkilölle on haasteellista. Sen sijaan eksplisiittinen tieto voidaan tallettaa ja siirtää helposti, sillä se on usein kirjalliseen muotoon puettua tietoa. Eksplisiittisen tiedon ilmaisutapoja ovat esimerkiksi eri kielet ja matemaattiset ilmaisut. Koska data ja informaatio voidaan esittää yksiselitteisesti jollakin kielellä, kuten tietokonekieli tai matematiikka, niin niitä voidaan helposti pitää eksplisiittisenä tietona. Tietämyksestä ja ymmärryksestä johtavasta viisaudesta suurin osa taas on hiljaista tietoa, joka karttuu kokemuksen myötä. Tietämystä on kuitenkin mahdollista joltain osin myös siirtää eksplisiittisenä ihmisten välillä. (Laihonen et al. 2013, s. 18–19)

2.1.2 Tiedolla johtaminen

Tietojohtaminen viittaa englanninkieliseen termiin Knowledge management ja suomen kielessä se toimii yläkäsitteenä, joka sisältää sekä tiedon johtamisen että tiedolla johtamisen (Seppänen 2023). Tiedon johtaminen viittaa organisaation kykyyn oppia ja uusiutua, luoda uutta tietoa sekä hallita tietovarantoja ja -virtoja. Tiedolla johtaminen sen sijaan viittaa toimintatapoihin, joilla jalostetaan ja hyödynnetään organisaation tietoa sen toiminnan johtamisessa ja se tähtää tietoperusteiseen päätöksentekoon eli päätöksentekoon, joka tapahtuu analysoidun tiedon perusteella (Laihonen et al. 2013, s.32; Finto 2024). Tietojohtaminen on johtamista siten, että edistetään organisaation kykyä luoda arvoa tiedolla ja osaamisella. Sillä pyritään organisaatiossa olevan tiedon hyödyntämisen lisäksi myös sen varmistamiseen, että organisaatiossa tai sen saatavilla on tulevaisuudessa tarvittava tieto. Toisinaan

tietojohtamisella voidaan joskus tarkoittaa tiedon johtamista tai tiedolla johtamista. (Finto, 2024) Tieteenalana tietojohtaminen pyrkii selvittämään, miten tiedolla voidaan luoda arvoa liiketoimintaan sekä minkälaisilla johtamiskäytänteillä ja -välineillä tietoperusteisia resursseja johdetaan tehokkaasti ja miten tiedon avulla johdetaan organisaatioita (Laihonen et al. 2013, s.7).

Tietojohtaminen on ilmiönä ajankohtainen niin yksityisellä kuin julkisella sektorillakin ja siihen liittyvät termit esiintyvät usein julkisessa keskustelussa sekä erilaisissa yhteiskunnan ja yksityisen sektorin strategioissa ja julkilausumissa (Helander et al. 2020). Tietojohtamisen taustalla on ajatus tiedon merkittävästä roolista organisaatioiden menestymisessä. Se on suhteellisen uusi johtamisen osa-alue ja Suomessa siitä alettiin puhua 1990-luvulla. Tieto- ja viestintäteknologian nopea kehittyminen on tarjonnut uudenlaisia mahdollisuuksia datan ja informaation varastointiin, analysointiin ja välittämiseen, mikä on ollut keskeistä alan kehittymiselle. (Laihonen et al. 2013, s.6)

Tiedolla johtaminen on systemaattinen toiminta- johtamismalli, joka vaatii sen eri osat ja niiden hyödyntämiseen liittyvän osaamisen. Tämä johtamismalli ja siihen liittyvä datan analysointi yhdistettynä hyvään substanssiosaamiseen ja relevanttiin kokemukseen tuovat parhaiten arvoa silloin kun sitä toteutetaan kattavasti läpi koko organisaation toiminnassa. (Saarteinen et al. 2018, s.4) Tiedolla johtamisen tarkoituksena on auttaa organisaatiota moninaisten tavoitteiden toteuttamisessa, jotka voivat kohdistua käytännössä työntekijän tavoitteiden toteutumisen tukemisesta koko organisaation tavoitteisiin ja toisaalta koskettaa kaikkia organisaation toimintoja (Listenmaa 2023, s.48). Perusta tiedolla johtamiselle muodostuukin tavoitteiden ja niitä tukevien kannustimien asettamisesta, joka tapahtuu johtamisen kautta. Johtamis- ja toimintatapa, joka perustuu todennettuun tietoon systemaattisesti läpi koko organisaation muodostaa myös vahvan tiedolla johtamisen organisaatiokulttuurin. Tällainen organisaatiokulttuuri systematisoi tiedon ja datan hyödyntämisen ja vahvistaa tiedon roolia päätöksentekoa niin strategisella, taktisella kuin operatiivisella tasolla. (Saarteinen et al. 2018, s.4)

2.2 Datalähtöisyys päätöksenteon tukena

2.2.1 Data-analytiikka

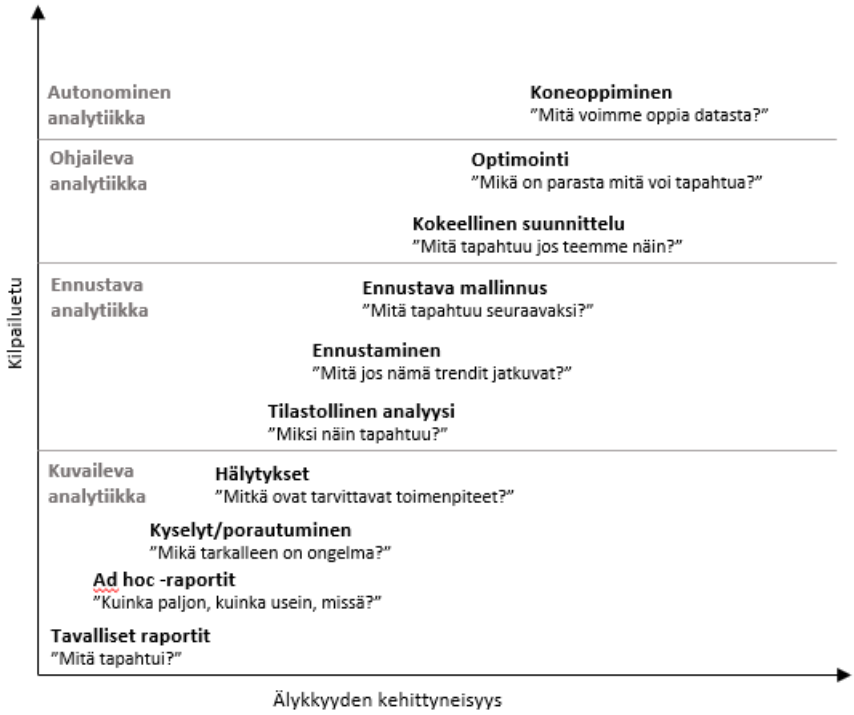
Data-analytiikka tai analytiikka voi olla kuvailevaa, ennakoivaa, ohjailevaa tai autonomista. Kuvaileva analytiikka (descriptive analytics), josta toisinaan käytetään termiä business intelligence eli BI tai suorituskykyraportointi, tarjoaa pääsyyn historiallisiin tietoihin. Sekä sisäistä että ulkoista dataa eri lähteistä käyttämällä se tarjoaa organisaatiolle mahdollisuuden luoda hälytyksiä, tutkia porautuvien kyselyiden avulla niin tavallista kuin Ad hoc -raportointia. Ennustava analytiikka (predictive analytics) käyttää kvantitatiivisia tekniikoita kuten segmentointi ja verkostoanalyysi sekä teknologioita, jotka ennustavat tulevaa historia tietoon pohjautuen. Ohjaileva analytiikka (prescriptive analytics) käyttää kvantitatiivisia tekniikoita, esimerkiksi mallinnusta, koneoppimista ja suosituskoneistoja optimaalisen käyttäytymisen ja toiminnan määrittämiseen. Autonominen analytiikka (autonomous analytics) käyttää tekoälyä tai kognitiivisia teknologioita, kuten koneoppimista, mallien luomiseen ja parantamiseen sekä datasta oppimiseen ilman inhimillisiä hypoteeseja ja huomattavasti vähemmällä analyytikoiden osallistumisella. (Davenport & Harris 2017)

Kukin analytiikan tyyppi tarjoaa vastauksia erilaisiin kysymyksiin, joiden avulla voidaan ymmärtää ja ennakoida ilmiöitä perusteellisemmin. Kuvaileva analytiikka keskittyy kysymyksiin, jotka käsittävät tapahtuneita asioita ja niiden yksityiskohtia, kuten *”Mitä tapahtui?”*, *”Kuinka paljon, kuinka usein, missä?”*, *”Mikä tarkalleen on ongelma?”* sekä *”Mitkä ovat tarvittavat toimenpiteet?”*. (Davenport & Harris 2017) Se esittää saatavan datan joko metrisissä muodoissa tai visuaalisesti kaavioiden avulla. On erittäin olennaista saada kuvaileva analytiikka oikein, jotta seuraavan tason analytiikka saadaan oikein. (Sharma et al. 2023, s.4)

Ennustava analytiikka sen sijaan pyrkii tunnistamaan syitä ja seurauksia sekä ennakoimaan tulevia trendejä. Se vastaa kysymyksiin *”Miksi näin tapahtuu?”*, *”Mitä jos nämä trendit jatkuvat?”* ja *”Mitä tapahtuu seuraavaksi?”*. (Davenport & Harris 2017) Nykyisen datakontekstiin perustuen sen tarkoituksena on yrittää kertoa mitä todennäköisesti tapahtuu. (Sharma et al. 2023, s.5)

Ohjaileva analytiikka arvioi eri toimintavaihtoehtojen vaikutuksia ja potentiaalisia lopputuloksia vastaamalla kysymyksiin *”Mitä tapahtuu, jos teemme näin?”* ja *”Mikä on parasta mitä voi tapahtua?”*. Sen painopiste on nimensä mukaisesti ohjata sopivimpaan toimintatapaan. Autonominen analytiikka pyrkii hyödyntämään dataa oppimisen välineenä ja se vastaa kysymykseen *”Mitä voimme oppia datasta?”*. (Davenport & Harris 2017; Sharma et al. 2023, s.5)

Kuvassa 2 on esitetty analytiikan neljä eri tasoa sekä niihin liittyvät kysymykset, joihin kukin taso pyrkii vastaamaan. Kuvasta huomataan, että potentiaalinen kilpailuetu kasvaa mitä kehittyneempää analytiikkaa on käytössä.



Kuva 2 Potentiaalinen kilpailuetu kasvaa kehittyneemmällä analytiikalla (Davenport & Harris 2017)

2.2.2 Business Intelligence - Liiketoimintatiedon hyödyntäminen

Liiketoimintatiedon hyödyntäminen tai liiketoimintatiedon hallinta ovat suomennoksia englanninkielisestä termistä BI eli Business Intelligence. Useissa suomenkielisissä teksteissä käytetään vielä englannin kielistä termiä, joten selvyyden vuoksi sitä tai sen lyhennettä käytetään myös tässä työssä.

BI:n määritelmälle on eri versioita, ja yksi niistä on Boris Evelsonin Forresterin esittelemä: ”*Business Intelligence on joukko menetelmiä, prosesseja, arkkitehtuureja ja teknologioita, jotka muuntavat raakadataa merkitykselliseksi ja hyödylliseksi tiedoksi, jota voidaan käyttää tehokkaampiin strategisiin, taktisiin ja operatiivisiin oivalluksiin sekä päätöksentekoihin.*” (van der Lans 2012). Business Intelligence on erittäin tärkeä kenttä organisaatioille kaikilla toimialoilla (Sabherwal & Becerra-Fernandez 2010). Se keskittyy organisaatioiden päätöksentekoprosessin tukemiseen ja mahdollisesti sen parantamiseen, toimittamalla oikeat tiedot oikeaan aikaan ja oikeassa muodossa. Siinä ei ole siis kyse päätöksen tekemistä, vaan sen tukemisesta. (van der Lans 2012) Toisen määritelmän mukaan vastaavasti BI-järjestelmät yhdistävät operatiiviset tiedot analyyttisiin työkaluihin, jotta monimutkainen ja kilpailukykyinen tieto voidaan esittää suunnittelijoille ja päätöksentekijöille. Tavoitteena on parantaa päätösprosessiin liittyvien panosten oikea-aikaisuutta ja laatua. BI:n avulla voidaan ymmärtää yrityksen kyvykkyydet kuten, tekniikan taso, trendit ja tulevaisuuden suunnat. (Negash 2004)

Päätöksenteko liittyy suoraan johtamiseen ja organisaation eri johtamistasoihin. Kuvassa 3 esitetty johtamispyramidi on suosittu tapa tunnistaa eri johtamisen tasoja. Ylimmän tason eli strategisen johtamisen painopiste on koko organisaation suorituksessa. Suurin osa tällä tasolla tehdyistä päätöksistä koskee koko organisaatiota ja niillä on pitkän aikavälin seurauksia sekä tavoitteita. Esimerkkikysymyksiä mihin tämän tason päätökset pyrkivät vastaamaan ovat muun muassa: “*Pitkäisikö meidän mennä uusille markkinoille?*”, “*Pitäisikö organisaatio olla tasoitettu?*” sekä “*Pitäisikö meidän ulkoistaa koko tietotekniikka?*”. Keskimmäinen taso eli taktinen johtaminen liittyy organisaation yksiköiden, kuten markkinointi ja myynti, suunnitteluun ja ohjaukseen. Yleensä tämän tason päätöksenteolla on vaikutusta lyhyellä tai keskipitkällä aikavälillä. Alin taso eli operatiivinen johtaminen käsittelee organisaation päivittäistä toimintaa. Esimerkkejä tällä johtamistasolla tehdyistä päätöksistä ovat: “*Onko tarpeeksi kuljettajia aikataulutettu kaikille toimitettavaksi oleville paketeille?*” ja “*Mihin*

toimenpiteisiin pitäisi nyt ryhtyä, kun vakava myöhästyminen on sattunut saapuvalle lentokoneelle, jossa on paljon matkustajia, joiden on suoritettava jatkolentoja?". (van der Lans 2012, s. 28–29)

Kuva 3 Johtamispyramidin tasot (mukaillen van der Lans, 2012 s. 28)

Päätöksenteon tueksi organisaatiot hankkivat BI-järjestelmiä ja tiedonanalysoinnin sovelluksia, jotka mahdollistavat perustellun päätöksen laajemman tiedon pohjalta. BI-järjestelmien tavoitteena on parantaa päätöksenteossa käytetyn tiedon laatua sen varastoinnin yksinkertaistamisen, tunnistamisen sekä analysoinnin seurauksena. Ne tarjoavat kattavan näkymän koko organisaatiosta, sallivat liiketoimintojen analysoinnin useista näkökulmista ja mahdollistavat nopean reagoinnin liiketoimintaympäristön muutoksiin. (Fink et al. 2017, s. 38–56)

Ralph Kimball ja Margy Ross kuvaavat kirjassaan (2013, s. 3–5) perusvaatimuksina, että BI-järjestelmän on:

- tehtävä tiedot helposti saatavaksi
- esitettävä tiedot johdonmukaisesti
- mukauduttava muutoksiin
- esitettävä tiedot oikea-aikaisesti
- oltava turvallinen linnake, joka suojaa tieto-omaisuutta

- toimittava arvovaltaisena ja luotettavana perustana paremmalle päätöksenteolle
- tultava hyväksytyksi liiketoimintayhteisön toimesta, pitääkseen sitä onnistuneena.

Vaikka jokainen näistä perusvaatimuksista on tärkeä, kaksi viimeistä ovat kriittisimmät ja valitettavasti usein unohdetuimmat (Kimball & Margy 2013, s. 3–5).

2.2.3 Perinteinen ja itsepalvelu BI

Perinteisesti BI:ssä on käytetty keskitetyttä datavarastoa ja IT-pohjaisia raportointiprosesseja. Modernimpi lähestymistapa on itsepalvelu BI (self-service BI), jonka avulla organisaation edustajat voivat käyttää ja analysoida dataa ilman laajaa IT-osaston osallisuutta. Se voi sopia paremmin organisaatioille, jotka etsivät enemmän ketteryyttä ja nopeampia oivalluksia tai niille, jotka haluavat edistää tietopohjaista kulttuuria ei-teknisten käyttäjien keskuudessa. Toisaalta perinteinen BI ratkaisu voi olla parempi organisaatioille, jotka tarvitsevat korkeatasoista datan ja raportoinnin hallintaa, joilla on monimutkaisia tietojen integrointitarpeita tai jotka vaativat korkean tason yhdenmukaisuutta ja turvallisuutta. Joka tapauksessa, sekä perinteisesti että itsepalveluna toteutettua BI:tä käytetään datan analysointiin ja tulkitsemiseen päätöksenteon tukena. (Atlan 2023)

Perinteisessä BI ympäristössä IT osasto kerää raakadataa eri lähteistä, puhdistaa sekä organisoi sen, jonka jälkeen tallentavat sen tietovarastoon. Analyytikot sitten käyttävät erilaisia työkaluja tehdäkseen kyselyitä, laatiakseen raportteja ja luodakseen oivalluksia tästä datasta. Yleisimpiä perinteiseen BI:hin liittyviä osia ovat tietovarasto, ETL-työkalut, OLAP, Raportointi- ja kyselyohjelmistot sekä datan koontinäytöt. Vaikka perinteiset BI-järjestelmät ovat olleet erittäin tehokkaita auttamaan yrityksiä käyttämään dataa päätöksenteossa, niin niille on tyypillisesti ominaista huomattava monimutkaisuus. Ne edellyttävät usein IT-ammattilaisten osallistumista tietojen hallintaan ja raporttien luomiseen, mikä voi johtaa viivästyksiin ja rajoittaa yrityskäyttäjien mahdollisuuksia saada oikea-aikaista ja olennaista tietoa. Monimutkaisuudesta huolimatta, Perinteisessä BI:ssä on etunsa, joista keskeisimmät ovat hallittu ympäristö, monimutkaisen datan käsittely, turvallisuus sekä aikataulutettu raportointi. (Atlan 2023)

Itsepalvelu BI:n on tarkoitus tuoda data ja oivallukset kaikkien organisaation jäsenten, eikä vain datatieteilijöiden ja analyytikoiden, saataville. Sen tarkoituksena on vähentää päätöksenteon viiveitä, joita usein ilmenee silloin kun päätöksentekijöiden täytyy luottaa siihen, että IT-henkilöstö luo raportit ja oivallukset datasta. Itsepalvelu BI:n avulla käyttäjät voivat tuoda itse dataa, luoda omia koontinäyttöjä ja suorittaa ad-hoc -analyysejä. Se tarjoaa joustavuutta personalisoitujen raporttien luomiseen ja datan perusteelliseen tutkimiseen. (Atlan 2023) Täten satunaiset käyttäjät pystyvät suorittamaan suurille määrille monipuolista dataa mukautettua analytiikkaa ja saada niistä käyttökelpoista tietoa ilman tarvetta ottaa mukaan BI-asiantuntijoita. Toisaalta tehokäyttäjät voivat suorittaa tehtävänsä sen avulla entistä helpommin ja nopeammin. (Alpar & Schulz 2016)

2.2.4 Datan visualisointi: tieto näkyväksi

Termi datan visualisointi viittaa ”tietokoneella tuettujen, vuorovaikutteisten, visuaalisten dataesitystapojen käyttöön kognition eli tiedon hankinnan ja käytön vahvistamiseksi” (Janvrin et al. 2014, s 44.). Datan visualisointi esittää tiedot kuvallisesti ja usein graafisesti mikä auttaa viestimään intuitiivisemmalla tavalla. Visualisointitekniikoiden käyttäminen datan yhdistämiseen ja sovittamiseen voi auttaa näkemään sen uudella tavalla ja saamaan siitä uusia oivalluksia. (Sherman 2014, s. 104) Hyvän visualisoinnin avulla voidaan nähdä trendejä, kuvioita ja poikkeavuuksia datasta. Yau:n (2013) mukaan paras visualisointi herättää autuuden hetken silloin kun näkee jotakin ensimmäistä kertaa ja tietää, että tieto on ollut aivan vieressä, mutta vain hieman piilossa. Joskus tämän hoitaa yksinkertainen diagrammi, mutta toisinaan visualisointi on kompleksisempi datan takia. Datan visualisointia pidetään yhä enenevissä määrin tehokkaana työkaluna sekä datan ymmärtämiseen että siitä kerättyjen oivallusten välittämiseen yleisölle (Kavitha 2022).

Datan visualisoinnin tavoitteena on helpottaa ymmärtämistä (Kirk 2019, s. 37). Aigner et al. (2023, s.4) mukaan tavoitteen saavuttamista voidaan mitata kolmen kriteerin avulla, jotka ovat ilmaisevuus, tehokkuus ja tarkoituksenmukaisuus. Hyvin abstraktilla tasolla voidaan eritellä visualisoinnin tekemisen kolmeksi perustavoiteeksi tutkiva analyysi, vahvistava analyysi ja analyysitulosten esittäminen. Tutkiva analyysi voidaan ymmärtää ohjaamattomana etsintänä, jossa annetusta datasta ei ole aikaisempia hypoteeseja. Tavoitteena on saada

käsitys datasta, aloittaa olennaisen tiedon poimiminen ja esittää hypoteeseja. Vahvistavassa analyysissä visualisointia käytetään todistamaan tai hylkäämään hypoteeseja, jotka voivat olla peräisin datan tutkimisesta tai dataan liittyvistä malleista. Siinä mielessä vahvistava analyysi on ohjatun etsinnän muoto. Lopulta kun faktat datasta on todettu niin esityksen visualisointien tavoitteena voi olla kommunikoida ja levittää analyysituloksia. (Aigner et al. 2023, s.4–5)

Kaaviotyypit

Datan visualisoimiseen käytetään eri kaavioita. Kaaviotyyppejä on monia erilaisia, joista yleisimmät ja yksinkertaisimmat ovat pylväs-, ympyrä- ja viivakaavioita. Riippuen siitä missä muodossa data on ja mitä visualisoinnilla halutaan esittää, on olennaista valita oikea kaaviotyyppi. Aina ei kuitenkaan ole yhtä oikeaa valintaa vaan eri vaihtoehdot voivat olla yhtä hyviä eri tavalla, kuvaten dataa eri näkökulmasta. (Unwin 2008, s.64)

Pylväskaavio on yksinkertaisesti graafinen tapa tehdä nopeasti laadullisia vertailuja arvojoukosta piirtämällä jokainen joukon arvo suorakaiteen muotoiseksi palkiksi, jonka pituus on verrannollinen sen edustamaan arvoon. Kaksi kertaa pidemmällä tangolla on kaksinkertainen arvo; palkki, joka on vain vähän pidempi kuin seuraava, edustaa vain hieman korkeampaa arvoa. (Wyatt & Johnsson 2009, s.36) Pylväskaaviot toimivat hyvin varsinkin vertailussa ja sijoitussuhteita kuvattaessa. Ne voivat olla esitetty vaakapalkein tai pystypalkein. Tyypillisesti jos sijoitussuhdetta korostetaan niin vaakapalkit saattavat toimia paremmin, koska ihmiset ovat tottuneet sijoituksiin, jotka menevät ylhäältä alaspäin (Ranganathan 2022, s. 7–8). Pylväskaavioita käytetään usein yksittäisten arvojen eikä niinkään yleistrendien korostamiseen. Skaalaus on katsottu joskus sen haittapuolena ja mikäli arvoasteikko ei ala nollasta, niin tulkinnat voivat olla harhaanjohtavia (Jääskeläinen et al. 2016, s. 15). Viivakaavio näyttää tiedot datapisteiden sarjana, jotka on yhdistetty toisiinsa viivasegmenttien avulla. Se on perusvisualisointi, joka on yleinen monilla aloilla ja erityisen tehokas näyttämään yleiskuvan koko tietojoukosta. (Jääskeläinen et al. 2016, s. 15). Ympyräkaaviot ovat hyviä näyttämään osuuksia, kun tarkastelussa on pieni määrä luokkia (Unwin 2008, s.64). Niitä on kritisoitu tehottomina, mutta myös Ranganathan (2022, s. 17) mukaan niitä voi käyttää tilanteissa missä kategorioiden määrä on pieni tai varsinkin jos niitä on kaksi. Neljän osan ympyräkaaviot ovat “vielä siedettäviä”, mutta siihen kun lisää luokkien määrää niin kaaviosta tulee yhä

lukukelvottomampi. Toinen tapaus missä piirakkakaavion käyttö soveltuu Rangathan mukaan, on jos halutaan näyttää, että luokkien ryhmän summa on pienempi kuin jäljellä olevien luokkien summa. Taulukossa 2 on kuvattu Goldenstone et al. (2015) esittämät päävisualisoinnit.

Taulukko 2 Päävisualisoinnit (mukaillen Goldstone et al. 2015)

Visualisoinnin nimi	**Kuvaus**
Taulukot	Järjestetty ruudukossa rivien ja sarakkeiden avulla. Ruudukon solut voivat sisältää geometrisiä, kielellisiä tai kuvallisia symboleja. Esim. tauluko ja matriisit.
Kaaviot	Kuvaa kvantitatiivista ja kvalitatiivista dataa käyttämättä hyvin määriteltyjä viitejärjestelmiä
Diagrammit	Piirtää kvantitatiiviset ja/tai kvalitatiiviset datamuuttujat hyvin määritellyssä viitejärjestelmässä, kuten koordinaatit vaaka- tai pystyakselilla.
Kartat	Näyttää datatietueet visuaalisesti, niiden spatiaalisten suhteiden mukaisesti ja näyttävät miten data jakautuu alueellisesti.
Verkostoasettelut	Käyttää solmuja datatietueiden edustamiseen ja yhdistämiseen linkityksin toisiinsa.

Asettelu

Graafiin voidaan luoda visuaalinen hierarkia tai painopisteitä lisäämättä lisäelementtejä, soveltamalla Gestaltin-periaatteita. Niiden mukaan havaintoon ei vaikuta vain elementit vaan myös konteksti. Graafiset elementit katsotaan osaksi ryhmää, kun ne ovat lähellä toisiaan (läheisyys), niillä on samanlainen väri tai varjostus (samankaltaisuus), ne ovat kohdistettu näkymätöntä viivaa tai käyrää pitkin (jatkuvuus) tai ne on sijoitettu niin, että ne näyttävät olevan suljetun muodon sisällä (sulkeminen). Visualisointienasettelussa voidaan hyödyntää Gestaltin periaatteita, jotta lopputuloksesta tulisi mahdollisimman intuitiivinen (Myatt & Johnson 2009, s. 24–25)

Kuvaajien tulisi olla juuri tarpeeksi isoja, jotta luettavuus toteutuu, muttei kuitenkaan paljoa tätä isompia. Luettavuus riippuu paljolti kuitenkin siitä, mitä muuta sisältöä samalla sivulla on esitettynä. Taulukoiden erottelu toisistaan on mahdollista kehysten avulla, mikä parantaa luettavuutta. Ei ole kuitenkaan kannattavaa käyttää liikaa tilaa kehyksille. (Unwin 2008, s. 68)

Väritys

Väritys on yksi tärkeimmistä huomioitavista asioista modernissa visualisoinnissa. Värivalintaa päätettäessä on harkittava useampaa asiaa. Ensiksi pitää muistaa, että viidellä prosentilla ihmisistä on jonkinlainen värisokeus, joista yleisimpänä punavihreävärisokeus. Toiseksi ihannetapauksessa värivalinnan tulisi myös välittyä tehokkaasti silloin kun visualisointi tulostetaan harmaasävyinä, koska artikkelit todennäköisesti luetaan joskus siinä muodossa. Tärkeimpänä tulee harkita esitettävän tiedon tyyppiä. Värivalinta tulisi riippua sen mukaan onko data kategorista, jatkuvaa ja onko siinä nollapistettä. (Hehman & Sally 2021) Myatt & Johnson (2009) nostavat tärkeimmäksi säännöksi värin valinnassa sen, ettei koskaan tee mitään mikä tekee lukemisen mahdottomaksi.

2.3 Datamallin kehittämisen keskeinen terminologia

2.3.1 Datan mallintaminen

Datan mallintamien on prosessi, jossa luodaan visuaalisia esityksiä datarakenteiden välisistä suhteista. Näiden esitysten avulla voidaan kuvata kokonainen tietojärjestelmä tai sen osia, ja ne välittävät tiedon siitä, miten datapisteet ja rakenteet liittyvät toisiinsa. (IBM 2024) Datamalliksi kutsutaan joukkoa taulukoita mitkä on linkitetty toisiinsa suhteiden avulla (Russo et al. 2020, s. 1). Datan mallinnuksen päätavoitteena on luoda jäsennelty esitys datasta, joka auttaa organisaatiota hallitsemaan, tallentamaan ja käyttämään sitä tehokkaasti. Mallinnuksen kriittisimpiin tarkoituksiin kuuluvat datan organisointi, selkeys, integrointi, laadunvarmistus, tehokas kysely, skaalautuvuus, päätöksen tukeminen, viestintä, tietokannan suunnittelu, tietojen hallinnointi, muutosten hallinta ja dokumentaatio. (Atlan 2023) Datan mallintamiseen ja tietovarastoiden luomiseen olennaisesti liittyy ETL-prosessi, millaiseen rakenteeseen taulukot on muodostettu sekä niiden väliset suhteet.

ETL-prosessi

ETL-prosessin kolme kirjainta tulee englanninkielisistä sanoista *Extract*, *Transform* ja *Load*, jotka ovat suomeksi käännettynä *poiminta*, *muuntaminen* ja *lataus*. Prosessissa data haetaan ja muunnetaan lähdejärjestelmästä siten, että se voidaan sijoittaa tietovarastoon (Rainardi 2008, s. 173). Goldfedder (2020, s. 76) mukaan puhtaimmillaan ETL-prosessissa data haetaan lähteestä suorittaen sille jokin toiminto sen muuttamiseksi tai manipuloimiseksi sen kulkua, jonka jälkeen kyseisen datan viimeisin versio työnnetään haluttuun kohteeseen. ETL-prosessia, jota toisinaan kutsutaan myös dataputkeksi, on erittäin hyödyllinen BI-projektien korkealaatuisen tiedon tuottamisessa (Sharma et al. 2023, s. 49). ETL on tärkeässä osassa tietovaraston ja datamallien kehittämisessä, jotta eri lähdejärjestelmien tiedot saadaan integroitua yhdenmukaiseen ja johdonmukaiseen muotoon (Minh et al. 2022, s. 498). ETL tapahtuu kerran, silloin kun tietovarasto täytetään ensimmäisen kerran. Lisäksi se tapahtuu aina sitten kun tietovarasto päivitetään. (Golfarelli & Rizzi 2009, s. 37)

Poiminnalla tarkoitetaan datan hakemista jo tunnistetuista lähteistä, eikä niinkään sen etsimistä. ETL-prosessilla on tarkoitus automatisoida datan poiminta ja sen jalostaminen. Datan poiminta tietovarastoon aloittaa varsinaisen ETL-prosessin (Kimball & Ross 2013, s. 19). Inkrementaalinen poiminta, mitä käytetään tietovaraston säännöllisessä päivittämisessä, ottaa haltuunsa lähdedataan tehdyt viimeisimmät muutokset. Se perustuu usein tietokannan ylläpitämään lokiin. Poiminnan prosessia voidaan suoraviivaistaa, jos operoitava data käsittää aikaleimanmerkinnän aina kun dataa muutetaan tai lisätään. Poiminta voi olla myös lähdepohjaista, jos sinne tehdyt toimet pystytään uudelleenkirjoittaa ilmoittaen asynkronisesti tehdyistä muutoksista tai jos käytössä oleva tietokanta voi toteuttaa asiaankuuluvan datan muutostapahtumiin liittyviä laukaisimia. (Golfarelli & Rizzi 2009, s. 39)

ETL-prosessin toisessa eli muuntaminen -vaiheessa tapahtuu tietojen muuntaminen. Goldfedder (2020, s. 78) kiteyttää, että muuntamisvaiheessa data muokataan, analysoiden sen eheys sekä vahvistaen tuotosdata. Tieteellisessä kirjallisuudessa puhdistuksen ja muuntamisen rajat ovat usein epäselvät terminologisesta näkökulmasta. Joskus ne eritellään omiksi vaiheiksi ja joskus ovat samaa vaihetta. Puhdistamisen on tarkoitus parantaa datan laatua, joten se on tärkeä tietojärjestelmän toiminnan kannalta. Yleisimpiä virheitä ja epäjohdonmukaisuuksia, jotka tekevät datasta “likaista” ja joita puhdistetaan ovat muun muassa datan kaksoiskappaleet, puuttuva data, odottamaton kenttien käyttö, mahdottomat tai väärät arvot,

erilaisista käytännöistä johtuvat epäjohdonmukaiset arvot yhdellä kokonaisuudella sekä epäjohdonmukaiset arvot kirjoitusvirheiden vuoksi. (Golfarelli & Rizzi 2009, s. 38–39)

ETL-prosessin viimeinen vaihe on lataus, missä tiedot ladataan tietovarastoon. Se voidaan tehdä kahdella tapaa päivittämällä. Ensimmäinen tapa on päivittää data siten, että se uudelleenkirjoitetaan kokonaan, jolloin vanha data korvataan. Tätä tapaa käytetään yleensä yhdessä staattisen poiminnan kanssa, kun tietovarastoon ladataan data ensimmäistä kertaa. Toinen päivitystapa on ladata vain ne lähdedatat, joihin on tehty muutoksia. Tämä suoritetaan yleensä ilman että olemassa olevaa dataa poistetaan tai muokataan. Usein tätä käytetään inkrementaalisen poiminnan kanssa yhdessä, kun tietovarastoa päivitetään säännöllisesti. (Golfarelli & Rizzi 2009, s. 41)

ETL-mahdollistaa instanssitason ja skeematason käsittelyn. Instanssitason käsittely pyrkii poistamaan itse olevia virheitä, kuten näppäily- tai kirjoitusvirheitä. Toisaalta skeematason käsittely käsittää tyypillisesti tietokannan muuntamisen uudeksi skeemaksi tai tietovarastoksi. (Kalita J. K. et al. 2024, s. 36)

Suhteet

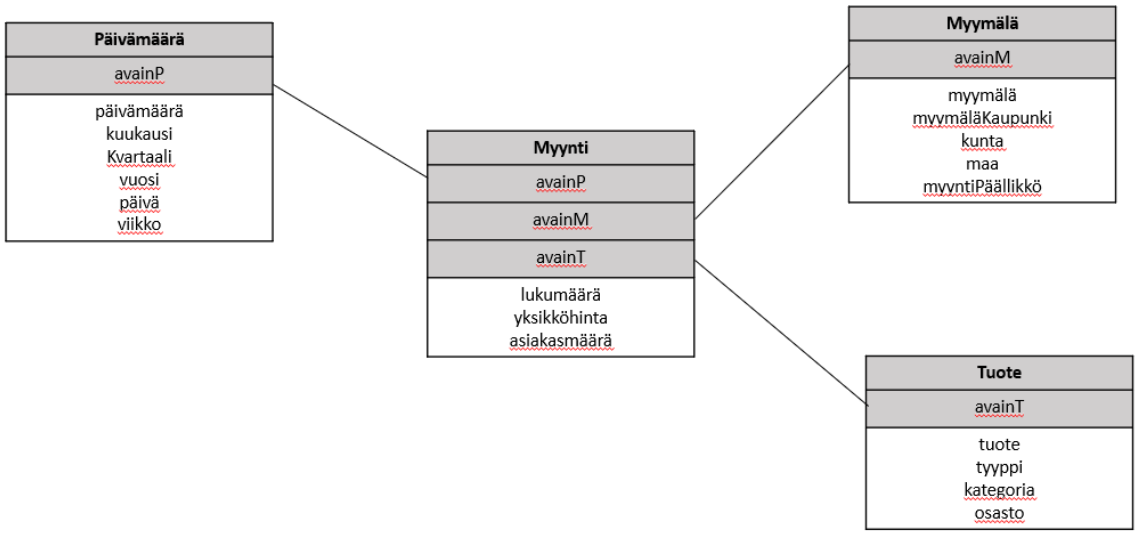
ETL-prosessin valmistuessa käyttövalmis data on ladattu taulukoina kohdejärjestelmään. Taulukko on joukko rivejä jaettuna sarakkeisiin mitkä sisältävät dataa. Jos datamallissa on useita taulukoita, ne voidaan linkittää toisiinsa suhteiden avulla. (Russo et al. 2020, s.2) Suhde yksinkertaisesti määrittelee, kuinka eri taulukot liittyvät toisiinsa, esimerkiksi asiakastietoja käsittävä taulukko voi olla liitettynä myyntitietoja käsittävään taulukkoon avainsarakkeen avulla. Suhteet voidaan rakentaa vain yhdelle sarakkeelle, ei monelle. (Knight et al. 2020, s. 52) Kun avainsarake on toisessa taulukossa, sitä kutsutaan vierasavaimeksi ja kun se on niin kutsutussa “kotitaulukossa” se on pääavain. Täten kaksi taulukkoa voidaan yhteen sovittaa liittämälle ne toisiinsa vierasavaimen ja pääavaimen sarakkeiden avulla. (Jolly 2023, s. 63)

Taulukoiden välisellä viivalla kuvataan aktiivista suhdetta. Suhteilla on myös suunta ja kardinaliteetti, mikä voi olla yksi-yhteen, yksi-moneen tai moni-moneen (Silverston 2001, s.12). Yleensä ne ovat yksi-moneen, mistä esimerkinä asiakastietoja käsittävästä taulukosta

(yksi -puoli) suhde menee myyntitietoja käsittävään taulukkoon (moni -puoli) eli yhdellä asiakkaalla voi olla monta myyntitapahtumaa. (Clark 2020, s. 80)

Tähtirakennemalli

Moniulotteinen mallinnus relaatiojärjestelmissä perustuu niin kutsuttuun Tähtirakennemalliin (Star Schema) ja sen muunnelmiin (Golfarelli & Rizzi 2009). Tähtirakennemalli koostuu dimensio- ja faktatauluista ja se saa nimensä sen graafisesta esityksestään, jossa faktataulukko muodostaa keskipisteen ja dimensiotaulukot lähtevät tästä keskustasta säteinä muodostaen yhdessä tähden muotoisen mallin (van der Lans 2012, s.41). Termi viittaa datamallin muotoon, mutta se on enemmän metaforinen kuin fyysinen. Siinä on yleensä vain yksi faktataulu, joka muodostaa keskustan ja jota ympäröi suhteilla linkitetyt dimensiotaulut. (Jolly 2023, s. 108) Kuvassa 4 on esitetty yksinkertainen esimerkki tähtirakennemallista, jossa faktatauluna keskellä on ”*Myynti*” ja dimensiotaulut ”*Päivämäärä*”, ”*Myymälä*” sekä ”*Tuote*” sen ympärillä.



Kuva 4 Esimerkki yksinkertaisesta tähtirakennemallista (mukaillen Golfarelli & Rizzi 2009, s.244)

Faktataulut sisältävät tapahtumatyyppistä dataa organisaation liiketoimintaprosesseista ja sen jokainen rivi vastaa yhtä mittaustapahtumaa. Yksi mallinnuksen perusperiaatteista on, että faktataulukon kaikkien mittarivien on oltava samalla tarkkuudella yksityiskohdiltaan. Tämä varmistaa sen, ettei mittauksia lasketa sopimattomasti kahteen kertaan. Faktataulun hyödyllisimmät arvot ovat numeerisia ja additiivisia eli yhteenlaskettavia, kuten esimerkiksi myynnin dollarimäärät. Arvot voivat kuitenkin joskus olla myös puoliadditiivisia tai jopa ei-additiivisia. Puoli-additiivisia arvoja, kuten tilisaldoja, ei voida summata aikadimensioon peilaten ja ei-additiivisia arvoja, kuten yksikköhintoja ei voida koskaan lisätä yhteen. Näistä arvoista voidaan laskea lukumääriä ja keskiarvoja. Faktataulut tulisi pyrkiä pitämään mahdollisimman niukkoina ja välttää tyhjien arvojen täyttämistä nolla-arvoilla, koska se ylikuormittaa useimmat faktataulut. Niukkuudesta huolimatta faktataulut muodostavat yleensä 90 prosenttia tai enemmän tähtirakennemallin käyttämästä kokonaistilasta. Kaikissa faktatauluissa on kaksi tai useampi viiteavainta, jotka muodostavat yhteyden dimensiotaulukoiden perusavaimiin. Esimerkiksi kuvassa 4 ”*Myynti*” -faktataulun viiteavain ”*avainP*” on ”*Päivämäärä*” -dimensiotaulun perusavain. Kun kaikki faktataulun avaimet vastaavat dimensiotaulujen perusavaimia niin taulukot täyttävät viite-eheyden. Faktataululla on yleensä oma perusavain, joka koostuu viiteavainten joukosta. Tätä kutsutaan yleensä yhdistelmäavaimeksi. (Kimbal et al. 2013, s. 10–12)

Dimensiotaulut sisältävät liiketoimintaprosessien mittaustapahtumaa liittyvän tekstikontekstin. Ne kuvaavat kysymyksiä ”*kuka, mitä, missä, milloin, miten ja miksi*” (Kimbal et al. 2013 s. 13–15). Niissä kukin elementti omaa vain yhden rivin, missä sen ominaisuudet on kuvattu tarkemmin. Datamalleissa on yleensä useita dimensiotauluja ja ihannetapauksessa ne ovat suoraan suhteessa faktatauluun, mutta ne voivat olla myös yhteydessä toiseen dimensiotauluun. (Jolly 2023, s. 102) Jokainen dimensiotaulu sisältää perusavaimena toimivan sarakkeen sekä joukon attribuutteja, jotka kuvaavat dimensiota. Melkein kaikki tähtirakenteet sisältävät päivämäärä dimension, koska tyypillisesti data mitä tarkastellaan kuten liiketoimintadata sisältää ajanhetken, jolloin yksittäinen tapahtuma on tapahtunut. (van der Lans 2012, s. 41–42).

2.3.2 Datamallin laajentaminen DAX:illa

DAX tulee sanoista Data Analysis eXpressions ja se on Microsoft Power BI:n, Microsoft Analysis Servicesin ja Microsoft Power Pivot for Excelin ohjelmointikieli. Se luotiin vuonna 2010 ja on sittemmin saavuttanut suosiota sekä Excelin Power Pivot -tietomallien luomisessa, että Business Intelligence -yhteisössä Power BI:n mallien rakentamisessa. DAX on erityisesti suunniteltu laskemaan liiketoimintakaavoja datamallissa ja sen ensisijainen tavoite on auttaa järjestämään, analysoimaan, ymmärtämään ja ehentämään dataa analysointia ja raportointia varten. (mukaillen Russo et al. 2020, s. 1 ja Seamark 2018, s. 2).

DAX on funktionaalinen ohjelmointikieli mikä koostuu lausekkeista, jotka periaatteessa ovat funktiokutsuja. Se ei sisällä monissa ohjelmistokielissä käytettyjä käsitteitä kuten luuppeja ja hyppyjä, vaan niiden sijaan kaikki on ilmaisua (Russo et al. 2020, s. 7).

DAX:in tärkeimmistä konsepteista ovat lasketut sarakkeet sekä mittarit. Vaikka ne muistuttavat toisiaan, koska tiettyjä laskemia voidaan tehdä kummalla tahansa, niin ne ovat todellisuudessa erilaisia. Laskettu sarake on tietomalliin lisätty uusi sarake, mutta tietolähteestä lataamisen sijaan se luodaan käyttämällä DAX-kaavaa. Se on aivan kuten mikä tahansa muu taulukon sarake ja sitä voidaan käyttää riveissä, sarakkeissa, suodattimissa tai matriisin arvoissa. Sen arvot ovat staattisia, kunnes data päivitetään. Tarvittaessa laskettua saraketta voidaan käyttää myös suhteiden määrittämiseen. Lasketulle sarakkeella määritetty DAX-lauseke toimii nykyisen rivin kontekstissa siinä taulukossa mihin laskettu sarake on luotu. Täten saman rivin muihin arvoihin voidaan viitata, mutta muiden rivien arvoihin ei ole suoraa pääsyä (Russo et al. 2020, s. 25; Knight et al. 2020, s. 98). Mittarit ovat sen sijaan hyödyllisiä silloin kun ei haluta laskea arvoja jokaiselle riville vaan halutaan koota arvot useilta taulukon riveiltä. Mittarien arvot eivät ole staattisia ja ne toimivat raportin sen hetkisessä suodatinkontekstissa, mikä tekee niistä dynaamisia ja suodatinkontekstin muutoksin mukautuvia. Kaikissa laskuissa ei voida saavuttaa oikeata tulosta lasketulla sarakkeella ja jos niissä on käytettävä usealta riviltä koottuja arvoja sen sijaan että toimittaisiin rivi kerrallaan, on luotava mittareita. Sekä mittarit että lasketut sarakkeet käyttävät DAX-lausekkeita, mutta ero on evaluaation kontekstissa. Mittarin evaluaatio tapahtuu visuaalisen elementin kontekstissa tai DAX-kyselyn yhteydessä, kun taas laskettu sarake rivitasolla siinä taulukossa mihin se kuuluu. Vaikka mittari on määriteltävä DAX-kielen vaatimuksien mukaisesti taulukkoon

niin se ei varsinaisesti kuulu sinne, joten sitä voidaan siirrellä taulukosta toiseen menettämättä sen toimivuutta (Russo et al. 2020, s. 26–29; Knight et al. 2020, s.108).

On olennaista harkita, milloin käyttää laskettua saraketta ja milloin mittaria. Joskus molemmat ovat vaihtoehtoja asian toteuttamisessa, mutta useimmiten laskentavaatimukset määrittävät valinnan. Laskettua saraketta tulee käyttää aina kun halutaan tehdä seuraavaa:

- asettaa laskennan tulokset osittajaan tai nähdä tulokset riveinä tai sarakkeina matriisissa tai pivot-taulukossa, tai käyttää laskettua saraketta suodatusehtona DAX-kyselyssä
- määrittää lauseke, joka on tiukasti sidottu nykyiseen riviin
- luokitella tekstiä tai numeroita esimerkiksi mittarin arvoalueiksi kuten 0-18, 18-25 ja niin edelleen.

Mittarin määrittäminen on pakollista aina kun halutaan näyttää käyttäjän valintoja heijastavia laskenta-arvoja ja arvot on esitettävä raportissa aggregaatteina kuten esimerkiksi laskea raporttivalikoiman voittoprosentti tai laskea yhden tuotteen suhdetta säilyttäen sekä vuosi että aluekohtainen suodatus. (Russo et al. 2020, s. 29–30)

Yleissääntönä toimii, että aina kun pystytään toteuttamaan laskenta kummalla tavalla tahansa, niin mittari on suositellumpi valinta ja laskettujen sarakkeiden käyttöä tulisi rajata niihin harvoihin tapauksiin, joissa niitä ehdottomasti tarvitaan. Excel-kokemusta omaavat käyttäjät suosivat yleensä mittareiden sijaan laskettuja sarakkeita, sillä ne muistuttavat suurelta osin tapaa millä laskutoimitukset suoritetaan Excelissä. Kuitenkin DAX:issa paras tapa laskea arvo on mittaria käyttämällä. (Russo et al. 2020, s. 30)

3 Tutkimus – Espoon kaupungin koulukuljetukset

Tämä luku käsittää tutkimuksen empiirisen osan. Luvun alkuun esitetään opinnäytetyön empiirisessä osassa käytetty tutkimusmetodologia ja menetelmät. Sen jälkeen tutkimuksen alustuksena on kohdeorganisaation esittely, missä kuvataan kohdeorganisaatio Espoon kaupunki, Logistiikan vastuuyksikkö, Koulukuljetusten järjestäminen Espoossa, henkilökuljetusdatan järjestelmätoimittaja sekä tiedoilla johtamisen malli Espoossa. Esittelyn jälkeen on Osa I mikä käsittää tutkimuksessa toteutetut haastattelut ja Osa II mikä käsittää Datamallin toteuttamisen.

Työ on kvalitatiivinen eli laadullinen tutkimus missä empiirinen osuus on toteutettu kahdessa osassa. Ensimmäisessä osassa haastatellaan valittujen kaupunkien edustajia ja kerätään aineistoa. Toisessa osassa rakennetaan datamalli Espoon henkilökuljetustiimin päätöksenteon tueksi koulukuljetusten osalta OSEMN-viitekehystä soveltaen.

Taulukko 3 Tutkimuksen empirian osat

Osa	**Tarkoitus**
1. Haastattelut	Kerätä vertailtavaksi aineistoa aiheesta: miten muut suomen kunnat hyödyntävät dataa koulukuljetuksien järjestämisen ja siihen liittyvän päätöksenteon tukena.
2. Datamalli	Luoda analyyttinen työkalu mitä Espoon kaupungin Logistiikkayksikkö voi hyödyntää työtehtäviensä ja päätöksenteon tukena ja edistää tiedolla johtamista.

3.1 Tutkimusmetodologia ja menetelmät

Tämä työ muodostettiin konstruktiivisen tapaustutkimuksen mukaisesti teoreettinen ja empiirinen aineisto yhdistämällä. Aineistona käytettiin kirjallisuuskatsausta sekä aineistoa, mikä saatiin haastatteluista ja kohdeorganisaation käytössä olevista järjestelmistä. Tapaus muodostui kohdeorganisaation tarpeesta kehittää tiedolla johtamista koulukuljetuksiin liittyen.

Konstruktiivinen tutkimus

Konstruktiivisen tutkimuksen tavoitteena on ratkaista käytännön ongelmia ja tuottaa samalla akateemisesti arvostettu teoreettinen tuotos. Ratkaisuna toimivat konstruktiot voivat olla prosesseja, käytäntöjä, työkaluja tai organisaatiokaavioita. Esimerkiksi nykyaikaisia talous-hallinnan järjestelmiä konstruoimalla voidaan ratkaista organisaatioiden todellisia ongelmia. Konstruktion tavoitteena on luoda uusi ja toimiva ratkaisu, joka vastaa selkeästi määriteltyyn ja aiempaan tietämykseen sidottuun ongelmaan. (Kasanen et al. 1991, s. 302–305; Pasian et al. 2015, s. 95) Konstruktiivisen tutkimuksen periaate on esitelty kuvassa 5.

Kuva 5 Konstruktiivisen tutkimuksen osat (Kasanen et al. 1991, 306)

Konstruktiivisessa tutkimuksessa erityisen tärkeää on, että molemmat osapuolet eli kohde-organisaatio ja tutkija ovat sitoutuneita kehitysprojektiin. Tyypillisesti käytännön edusta-jien edun mukaista on antaa tutkijalle käyttöön ja tutkimuksen raaka-aineeksi kaikki merki-tyksellinen aineisto ja paras osaaminen. (Lukka 2001)

Haastattelu

Tämän työn empiirisen osuuden ensimmäinen osa käsittää tutkimuksen haastattelut. Haas-tattelu on käytetyimpiä tiedonkeruumuotoja. Erityisesti vapaamuotoisten tai vähän struktu-roitujen haastattelumenetelmien käyttö on lisääntynyt. (Hirsjärvi & Hurme, 2022) Tämän työn haastattelut toteutetaan puolistrukturoituina haastatteluina, missä haastattelulle on mie-titty teemat, mutta niiden lisäksi on valmisteltu tarkkoja kysymyksiä, jotka kaikki esitetään haastateltaville. Puolistrukturoitu haastattelu sopii tilanteisiin, joissa on päätetty haluttavan tietoa juuri tietyistä asioista, eikä haastateltaville näin ollen haluta tai ole tarpeellista antaa kovin suuria vapauksia haastattelutilanteessa (KvaliMOTV 2022). Haastatteluissa

keskitytään keräämään vertailtavaksi asiaankuuluvia todisteita tutkittavasta aiheesta. Siten pystytään havaitsemaan ja ymmärtämään mahdollisia yhtäläisyyksiä, eroja, vahvuuksia ja haasteita haastateltavien kaupunkien välillä. Haastatteluista saatu aineisto vastaa työn toiseen tutkimuskysymykseen.

OSEMN-viitekehys

Tämän työn empiirisen osuuden toinen osa eli datamalli rakennetaan OSEMN-mallia (kuva 6) soveltaen. OSEMN on akronyymi, missä kukin kirjain tarkoittaa yhtä datamallin vaihetta. Se tulee sanoista Obtain (Hanki), Scrub (Puhdista), Explore (Tutki), Model (Mallinna) ja iNterpret (Tulkitse). (Hotz 2023)

Obtain (Hanki) on OSEMN-mallin ensimmäinen vaihe ja siinä keskitytään datan hankintaan. Data voi olla peräisin jo olemassa olevasta, vastikään hankitusta, vastikään kysellystä tai poimitusta datasta (Nantasenamat 2020). Tässä vaiheessa tulee ymmärtää mistä tarvittava data löydetään. Jos dataan on jo pääsy niin tee sille kysely. Kun sinulla ei ole pääsyä datoihin niin pyydä pääsy niihin. Mikäli data on ostettavissa, tee kustannus-hyötyanalyysi ja osta se, jos se katsotaan tarkoituksenmukaiseksi. Luo järjestelmä datan keräämistä varten siinä tapauksessa, jos dataa voidaan itse luoda. (Hotz 2023)

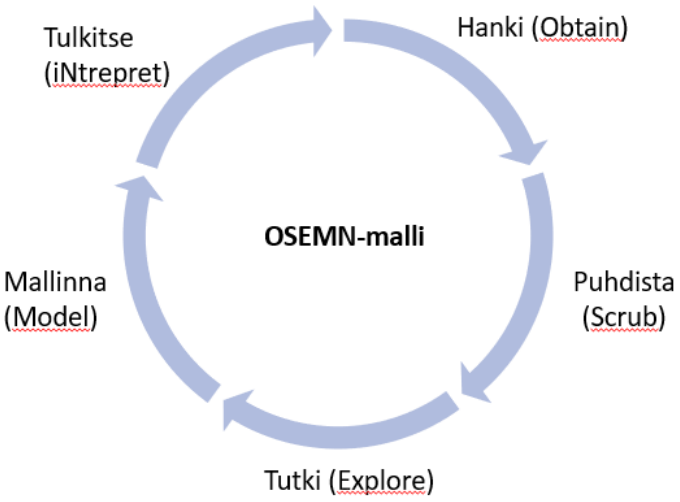
Scrub (Puhdista) -vaiheessa dataa lähinnä puhdistetaan, jotta raakadata saadaan muutettua käytettävään muotoon. Tätä vaihetta pidetään kaikkein aikaa vievimpänä, koska siihen sisältyy puuttuvien datojen käsittelyä sekä datan esikäsittely mahdollisimman virheettömäksi ja yhdenmukaiseksi (Nantasenamat 2020). Puhdistamisen vaiheet vaihtelevat projektin kokonaisuuden ja datasetin mukaan, mutta yleisiä vaiheita ovat datan laadun arviointi, epäolennaisen datan ja kaksoiskappaleiden poistaminen, puuttuvien arvojen lisääminen, eri datasettien yhdistäminen, datan lataaminen uuteen kohdepaikkaan sekä tehtyjen muunnosten dokumentointi (Hotz 2023).

Explore (Tutki) -vaihe on pohjimmiltaan kartoittava data-analyysi, missä saadaan käsitys datasta, jotta pystytään selvittää, miten voidaan toimia ja mitä alueita voidaan tutkia mallinnusvaiheessa. Tässä ei vielä testata hypoteeseja tai tehdä ennusteita vaan pikemminkin tutkitaan dataa eri tekniikoiden, kuten kuvailevien tilastojen ja visualisointien, avulla. Tavallisesti tässä vaiheessa ei kuitenkaan pyritä ymmärtämään jokaista yksityiskohtaa, vaan

saamaan hyvä käsitys, ennen kuin siirrytään seuraavaan vaiheeseen. (Nantasenamat 2020; Hotz 2023)

Model (Mallinna) -vaiheessa pyritään hyödyntämään koneoppimisen algoritmeja ja ymmärtämään dataa siten, että siitä saadaan hyödyllisiä oivalluksia, jotka ovat välttämättömiä päätöksenteossa. (Nantasenamat 2020)

iNterpret (Tulkitse) -vaiheessa saadaan oikeasti hyödynnettyä dataa, kun selvitetään, miten kaikkien rakennettujen mallien tuloksia voidaan yksinkertaistaa ja tiivistää. Tämä käsittää mielekkäiden johtopäätösten ja käyttökelpoisten oivallusten tekemistä, joiden avulla voidaan selvittää seuraavat tarvittavat toimenpiteet. Pohdittavana on niitä kysymyksiä, joiden perusteella koko prosessi ylipäätänsä aloitettiin, sekä sitä toiminnallista arvoa mitä datan tutkimisen ja mallintamisen prosessilla saavutetaan. Koko prosessin tarkoitus on tarjota näkemystä organisaatiolle tai sidosryhmille. Jos tätä ei saavuteta, niin hankkeen toteuttaminen oli lähinnä ajan ja resurssien tuhlausta. (Nantasenamat 2020; Hotz 2023)



Kuva 6 OSEMN-malli

3.2 Kohdeorganisaatio

Espoo on Suomen toiseksi suurin kaupunki yli 305 000 asukkaalla (Tilastokeskus 2023). Sen erityispiirre on viiteen eri keskukseen tukeutuva kaupunkirakenne, jossa yhdistyvät urbaani ja pientalovaltainen luonnonläheinen asuminen. Kaupungin henkilöstöä Espoossa on noin 10 500 (Espoon kaupunki 2023a&f).

Espoon strategiana on vuodesta 2013 lähtien joka valtuustokaudelle neljän vuoden välein päivitettävä Espoo-tarina. Siihen liittyvät poikkihallinnolliset kehitysohjelmat ja valtuustokauden tavoitteet ohjaa kaupungin toimintaa entistä paremmin ja selkeämmin yhteisten tavoitteiden mukaisesti. Nykyinen valtuustokausi on 2021–2025. (Espoon kaupunki 2023a)

Konsernihallinto ja toimialat muodostavat Espoon kaupungin hallinnollisen organisaation. Niiden toimintaa ohjaa ja valvoo luottamuselimistä koostuva päätöksenteko-organisaatio: valtuusto, kaupunginhallitus sekä lauta- ja johtokunnat. (Espoon kaupunki 2023b). Ylintä päätösvaltaa käyttää Valtuusto, jonka 75 jäsentä valitaan joka valtuustokaudelle vaaleilla kaupungin asukkaiden toimesta. (Espoon kaupunki 2023c)

3.2.1 Logistiikka (vastuuyksikkö) ja henkilökuljetukset

Konsernihallinnon Hallinto ja kehittäminen -yksikköön kuuluva Espoon kaupungin Logistiikka -vastuuyksikkö vastaa ja huolehtii kaupungin henkilökuljetuksista, joihin kuuluvat Sote-uudistuksen voimaanastuessa 1.1.2023 alkaen varhaiskasvatuksen kuljetukset ja koulukuljetukset. Logistiikka vastaa henkilökuljetuksiin liittyvästä sopimusseurannasta ja valvoo kuljetuspalvelun tuottamisesta yhteistyössä Kasvun ja oppimisen toimialan kanssa. (LOG Organisaatio 2023) Kuten taulukosta 1 huomataan, varhaiskasvatukset ovat kuljetusmäärällisesti ja kustannuksellisesti vain murto-osa koulukuljetuksista. Koska koulukuljetukset ovat määrällisesti merkittävämpi kokonaisuus ja kuten aikaisemmin kappaleessa 1.2 on mainittu, tämä tutkimus on rajattu koskemaan vain koulukuljetuksia.

Taulukko 4 Henkilökuljetusten kuljetusmäärät ja kustannukset vuodessa (LOG Organisaatio 2023)

Varhaiskasvatus	**Koulukuljetukset**
9 800 kuljetusta	310 000 kuljetusta
150 000 euroa	5,5 miljoonaa euroa

3.2.2 Koulukuljetukset Espoossa

Espoon kaupungin osoittamaa koulua käyvälle oppilaalle voi hakea koulumatkaetuutta yhden hänen väestörekisteriinsä merkityn Espoon osoitteen ja koulun välille. Etuuden voi saada, jos matka todetaan liian pitkäksi, vaikeaksi, rasittavaksi tai vaaralliseksi, huomioon ottaen oppilaan luokka-aste ja muut olosuhteet, kuten terveydelliset syyt. Maksuttoman joukkoliikenteen matkakortin saa, jos koulumatka on 1.–2. luokan oppilaalla yli kolme kilometriä ja 3.–10. luokan oppilaalla yli viisi kilometriä. Koulumatkaetuus myönnetään ensisijaisesti matkakorttina, mikäli etuutta haetaan koulumatkan vaikeuden, rasittavuuden ja turvattomuuden takia. (Koulukuljetusopas 2019)

Kuljetusyhtiön taksilla järjestetty koulukuljetus voidaan myöntää sellaisen lausunnon perusteella, jossa arvioidaan, kuinka oppilas selviytyisi koulumatkasta kävellen tai julkisia kulkuneuvoja käyttäen. Lausunnoksi kelpaa enintään kuusi kuukautta vanha lääkärin- tai psykologinlausunto, koulun oppimisen tuen ryhmän lausunto tai pedagoginen selvitys. Jokainen uusi koulumatka arvioidaan erikseen kaupungin toimesta. Jos oppilaan terveydentila, kuten vakava vamma tai sairaus, estää itsenäisen liikkumisen ja oppilas tarvitsee koulukuljetuksen useamman vuosiluokan ajaksi, koulukuljetus voidaan myöntää ainoastaan lääkärin lausunnon perusteella. (Koulukuljetusopas 2019)

Espoon koulukuljetusten osapuolena on useampi taho ja ne on kuvattu taulukossa 5.

Taulukko 5 Koulukuljetusten osapuolet

Osapuoli	**Tehtävät ja vastuut**
Kasvun ja oppimisen lautakunta sekä Svenska rum -lautakunta	Määrittää koulukuljetusten ja matka-avustusten myöntämistä koskevat linjaukset
Kasvun ja oppimisen toimiala (KOTO)	Päättää koulukuljetusetuuden myöntämisestä oppilaille.
Logistiikka	Vastaa koulukuljetuskokonaisuuden toimimisesta.
Koulu	Huolehtii omalta osaltaan siitä, että oppilaan siirtyminen koulusta kuljetukseen ja kuljetuksesta kouluun voi tapahtua sovittuna aikana ja turvallisesti. Lisäksi ilmoittaa aikatauluista ja muutoksista Logistiikkaan tai tekee ne suoraan kuljetustenhallintajärjestelmään.
Huoltaja	Hakee koulukuljetusetuutta oppilaalle ja on velvollinen ilmoittamaan oppilaan tietojen muutoksista. Huoltaja opastaa oppilasta noudattamaan turvallisen ja sujuvan kuljetuksen sääntöjä.
Kuljetusyritys ja kuljettaja	Kuljetusyritys ja kuljettaja huolehtivat matkojen reitityksestä sekä kuljetusten toteuttamisesta turvallisesti sovitussa aikataulussa.
Espoo-info (Asiakaspalvelu)	Ottaa vastaan aikataulumuutokset ja reklamaatiot
Järjestelmätoimittaja	Vastaa kuljetusten hallintaan liittyvien järjestelmien toimimisesta kaupungin ja kuljetusyhtiön päässä.

3.2.3 Järjestelmätoimittajan kuvaus

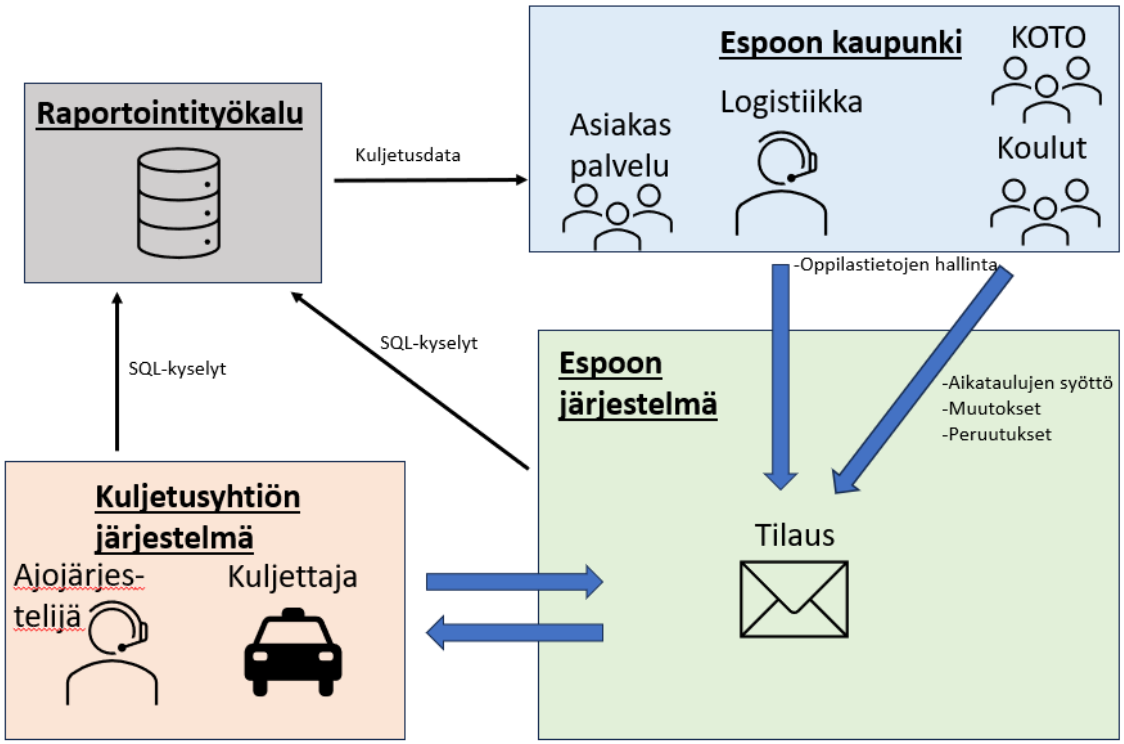
Järjestelmätoimittaja on kotimainen ensisijaisesti kuljetusten reitittämiseen erikoistunut yritys, joka tuottaa ohjelmistoa henkilökuljetusten järjestämiseen. Henkilökuljetusten suuri haaste on tuoda kuljetuspalvelun moninaiset osat yhden palvelukokonaisuuden alle. Kuljetuksiin liittyviä osapuolia on mm. kuljetettavat ja heidän huoltajansa, kuljetusyhtiöt ja kuljettajat, reittisuunnittelijat, koulut sekä kunta. Kaikilla osapuolilla on erilaiset intressit, tarpeet ja vaatimukset, jotka täyttämällä saavutetaan sujuva kuljetuspalvelu.

Espoon koulukuljetusten kokonaisuus järjestelmätoimittajan osalta koostuu kolmesta palvelusta: Espoon järjestelmä, Kuljetusyhtiöiden järjestelmä sekä Raportointityökalu.

Espoon järjestelmä = Tilausten ja matkustajatietojen hallintajärjestelmä on henkilökuljetusten hallintaan räätälöity selainpohjaisena tarjottava palvelu, mitä käyttää Logistiikka tiimi sekä koulut. Se mahdollistaa toistuvan ja vakiomuotoisen henkilöliikenteen aikataulujen hallinnan ja viikkoaikataulujen toimittamisen kuljetuksia hoitaville yhtiöille digitaalisesti. Järjestelmän avulla voidaan tehdä tilauksia, muutoksia tai peruutuksia rajapintojen yli suoraan kuljetusyhtiön reititysjärjestelmään.

Kuljetusyhtiön järjestelmä = Kuljetusten reitittämisen hallintajärjestelmät ja sen tietokannat. Palvelu on tarkoitettu suoraohjaukseen, henkilökuljetussuunnitteluun ja kuljetusten kokonaishallintaan, joka digitalisoi henkilökuljetuksen kuljetustilauksesta ajosuoritteeseen.

Raportointityökalu = Erillinen raportointityökalu, joka mahdollistaa Espoon järjestelmän tietokannasta tai siihen kytketyistä Kuljetusyhtiöiden tietokannoista tehtäviä kyselyitä. SQL-kyseluiden avulla, Logistiikkaa saa haettua dataa kuljetuksista raportointia varten. Toimii kokonaisuudessaan järjestelmätoimittajan palvelimella, jota kautta voidaan rakentaa kustomoituja raportteja Tilaajan (Espoon kaupungin) toimesta.



Kuva 7 Järjestelmätoimittajan tuottaman palvelun kuvaus Espoon kontekstissa

3.2.4 Tietojohtaminen, Excel- ja Power BI-raportointi Espoossa

Espoon kaupungissa tietojohtamisella tarkoitetaan johtamista siten, että edistetään organisaation kykyä luoda arvoa tiedolla ja osaamisella. Tietojohtamisella pyritään paitsi organisaatiossa olevan ja organisaation saatavissa olevan tiedon hyödyntämiseen myös sen varmistamiseen, että organisaatiossa tai sen saatavissa on tulevaisuudessa tarvittava tieto.

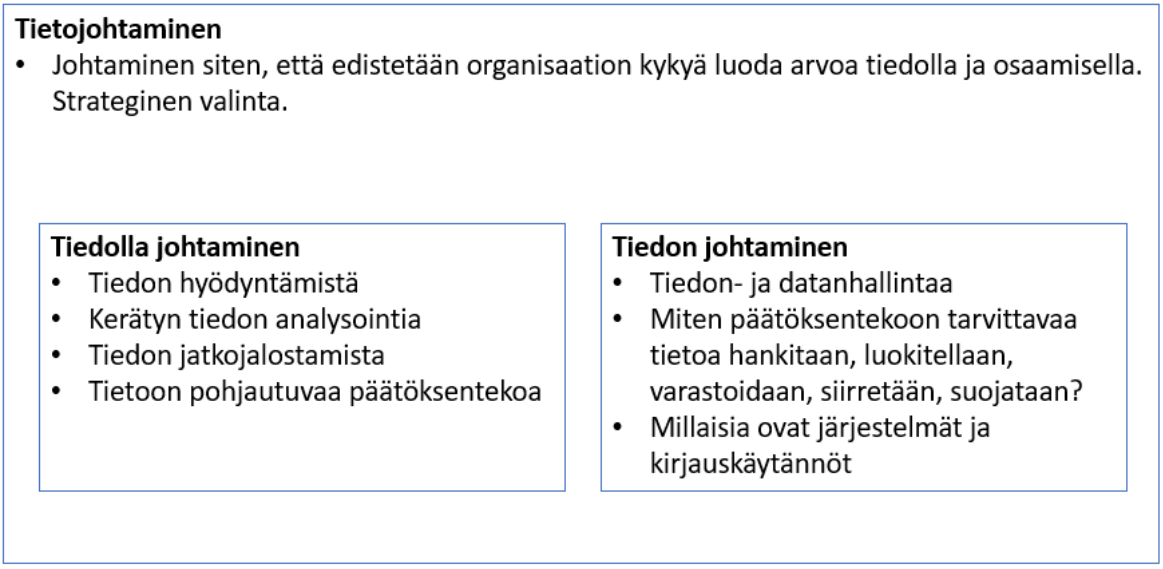
Tietojohtaminen on organisaation strateginen valinta. Espoo-tarinan mukaan organisaatiossa kehitetään tiedolla, data-analyysillä ja vaikuttavuudella johtamista. Espoo kumppaneineen on aktiivinen digitalisaation, robotiikan, tekoälyn ja muun teknologian hyödyntäjä läpi toimintaprosessien.

Tietojohtaminen jaetaan tyypillisesti kahteen osa-alueeseen:

1. Tiedon hyödyntäminen tai tiedolla johtaminen, jolla tarkoitetaan kerätyn tiedon analysointia ja jatkojalostamista siten, että voidaan tehdä tietoon pohjautuvaa päätöksentekoa.

2. Tiedonhallinta tai tiedon johtaminen, joka tarkoittaa niitä toimia, joilla tavoitellaan sitä, että käytettävissä on hyödyntämiskelpoista tietoa ja että tietoja hallitaan turvallisesti sekä lain ja muiden vaatimusten mukaisesti tiedon elinkaaren ajan.

Tietojohtamisen kehittäminen on yksi digitalisaation keskeinen osa-alue. Muita digitalisaation myötä kehitettäviä osa-alueita ovat muun muassa digitaaliset palvelut ja toiminnanohjaus. (Espoon kaupunki, 2023d)



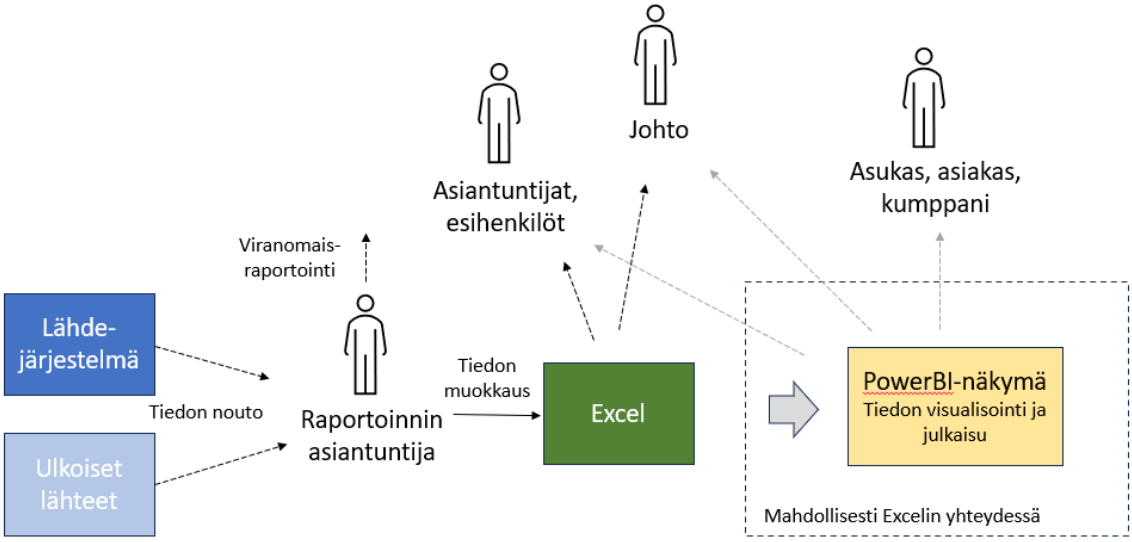
Kuva 8 Tietojohtaminen Espoon kaupungilla (Espoon kaupunki, 2023d)

Excel -raportointi Espoossa

Useissa tapauksissa Excel -taulukkolaskentaohjelman käyttö on raportoinnin lähtökohtana. Aluksi tiedot noudetaan manuaalisesti Excel -taulukkoon lähteestä tai useammasta lähteestä, jossa tieto jatkomuokataan siten, että taulukosta muodostuu raportti ja mahdollisesti myös sen visualisointi. Jatkomuokkaus voi olla manuaalista tai osin esimerkiksi makroilla automatisoitua. Tietojen tuonti voi olla osittain automatisoitua, esimerkiksi skripteillä, tai manuaalista tietojen hakua lähdejärjestelmien käyttöliittymistä, jotka usein ovat valmiiksi Excel- tai CSV-taulukkona.

Tähän Excel -raportoinnin kategoriaan katsotaan kuuluvan myös tilanne, jossa Power BI-työvälineellä visualisoidaan ja julkaistaan yksittäisessä Excel-työkirjassa ylläpidettyä tietoa. Excelin käytön etuna on, että se soveltuu joustavasti muuntuviin tietotarpeisiin, jossa laskentaa tehdään kertaluontoisesti tai hyvin paikallisesti. Sen sijaan haasteena on muun muassa versiointi ja ylläpidettävyys, sillä työkirjoista muodostuu helposti useita kopioita, nimeämiskäytännöt ovat vaihtelevia, tahattomia muokkauksia voi olla haasteellista estää etenkin yhteiskäytössä ja inhimillisen virheen mahdollisuudet korostuvat pyrittäessä toistamaan manuaalista prosessia samansisältöisenä. Dataa voi olla haasteellista pitää määrämuotoisena varsinkin jaetuissa työkirjoissa, jolloin esimerkiksi vertailtavuus vuosien välillä on haasteellisempaa.

Excelin käytön riittävän hyvä osaaminen on keskeinen menestystekijä Excel-taulukoiden hyödyntämisessä. Hyvin suunniteltu ja dokumentoitu työkirjan rakenne auttaa tuottamaan lisäarvoa, sujuvoittamaan työtä, välttämään virheitä ja mahdollistaa työkirjojen yhteiskäytön. (Espoon kaupunki, 2023e) Kuvassa 9 on kuvattu tavanomainen manuaalinen Excel-raportointi Espoon kaupungilla.

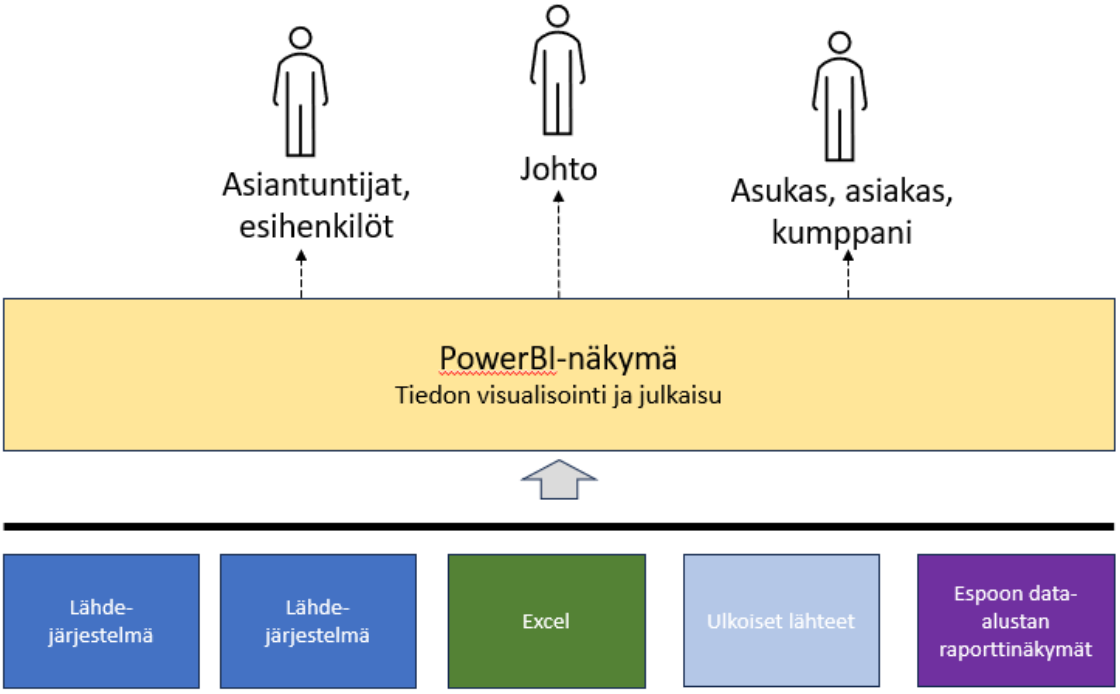


Kuva 9 Tavanomainen manuaalinen Excel- raportointi Espoon kaupungilla

Power BI –raportointi Espoossa

Microsoft Power BI on edistynyt raportoinnin ja tiedon visualisoinnin väline, jonka perusversio on käytettävissä kaikilla Espoon hallinnon työntekijöillä. Väline pystyy noutamaan tietoa monenlaisista tietolähteistä kuten järjestelmien tarjoamista API-rajapinnoista sekä Excel- ja CSV-taulukoista. Eri tietolähteistä noudettavia tietoja voidaan yhdistellä sillä edellytyksellä, että tiedot ovat yhdistämiskelpoisia, eli kun eri lähteistä tulevien tietojen avainkentät ovat yhteensopivia. Power BI-näkymiä voidaan julkaista sisäiseen tai ulkoiseen käyttöön Pro-lisenssillä.

Espoon tapauksessa Power BI:tä hyödynnetään yksikkötasolla käyttämällä tietolähteenä yhtä tai useampaa omaa lähdejärjestelmää. Yhteys muodostetaan lähdejärjestelmän tarjoamiin API-rajapintoihin tai lähdejärjestelmän tietokantaan, mitä suositellaan kuitenkin välttämään. (Espoon kaupunki, 2023e) Power BI -näkymien käyttö Espoon kaupungilla on nähtävissä kuvassa 10.



Kuva 10 Power BI-näkymien käyttö Espoon kaupungin raportoinnissa

3.3 Osa I: Tutkimuksen haastattelut

Tämän tutkimuksen kappale 3.3 käsittää työn empiirisen osuuden ensimmäisen osan. Osa käsittää teemahaastattelut, minkä avulla aineisto kerättiin tutkimusta varten. Haastattelut toteutettiin loka-marraskuussa 2023.

3.3.1 Haastattelujen toteutus

Tässä työssä haastattelut toteutettiin puolistrukturoituina teemahaastatteluina, minkä tueksi tehtiin haastattelurunko ohjaamaan haastattelua. Haastattelujen tavoitteena oli vertailla valittujen kaupunkien tapoja hyödyntää dataa koulukuljetusten järjestämisessä ja tunnistaa hyviä käytänteitä, joita voitaisiin ottaa käyttöön Espoon kaupungissa. Haastatteluista saaduilla tuloksilla onkin tarkoitus vastata toiseen tutkimuskysymykseen: *”Miten muut kunnat/kaupungit hyödyntävät dataa koulukuljetusten hallinnoinnissa ja mitkä ovat hyviä käytäntöjä mitä Espoon kaupunki voisi ottaa käyttöön toiminnassaan?”*.

Taulukko 6 Haastattelujen toteutuksen vaiheet ja aikataulu

Vaihe	**Ajankohta**
1. Haastattelurungon luominen	syyskuu 2023
2. Haastattelurungon läpikäynti ohjaajien kanssa ja täsmentäminen	syys-lokakuu 2023
3. Yhteydenotot haastateltaviin ja kutsujen lähettäminen	lokakuu 2023
4. Haastattelut ja litterointi	loka-marraskuu 2023
5. Haastattelujen analyysi	marras-joulukuu 2023
6. Tulosten kirjoittaminen	joulukuu 2023

Taulukon 6 mukaisesti haastattelut toteutettiin kuudessa vaiheessa ja niiden ajankohdat sijoittuivat syksylle sekä talvelle 2023. Ensimmäinen haastattelurunko ja sen kysymykset luotiin syyskuussa ja se katsottiin kohdeorganisaation yksikön ohjaajan kanssa läpi sekä lähetettiin yliopiston ohjaajalle kommentoitavaksi. Kommenttien perusteella haastattelu runkoa täsmennettiin ja lähetettiin toistamiseen ohjaajille kommentoitavaksi, jonka jälkeen haastattelurunko hyväksyttiin käyttöönotettavaksi haastatteluja varten. Lokakuussa haastateltaviin oltiin yhteydessä aluksi puhelimitse, missä esiteltiin haastattelun aihe, kysyttiin suostumusta haastatteluun sekä sovittiin alustava haastatteluajankohta. Puhelun jälkeen varsinainen haastattelukutsu sekä haastattelurunko lähetettiin sähköpostilla haastateltavalle. Parin haastattelun ajankohtaa jouduttiin siirtämään sairastumisen sekä yllättävien työkiireiden takia. Kaikki haastattelut toteutettiin loka-marraskuussa ja samalla niistä luotiin videotallenne ja litterointi aineiston keräämistä varten. Tämän jälkeen, marras-joulukuun aikana, aineisto koottiin yhteen ja analysoitiin minkä pohjalta kirjoitettiin haastatteluista saadut tulokset.

Haastatteluissa käytetty haastattelurunko koostui kolmesta teemasta sekä lisäyksistä haastattelun lopuksi. Haastatteluihin valitut teemat niitä tukevin kysymyksin olivat:

1. Johdanto ja koulukuljetuksiin tutustuminen
2. Datan hyödyntäminen, analytiikka ja raportointi organisaatiossa
3. Kehittäminen ja tulevaisuuden näkymät.

Haastattelun lopun lisäykset olivat vapaaehtoiset ja tarkoitettu sitä varten, mikäli jotakin jäi mainitsematta teemojen kysymysten yhteydessä. Pääkysymyksiä oli yhteensä 23 joiden tukena oli tarkentavia alakysymyksiä. Kysymysten tarkoitus oli tukea haastattelun kulkua, ohjaamatta kuitenkaan liikaa haastateltavaa. Haastattelurunko, mikä on esitetty liitteessä 1,

lähetettiin sähköpostilla ennakkoon haastateltaville 1–2 viikkoa ennen haastattelua, jotta he pystyivät tutustumaan kysymyksiin ja valmistelemaan vastauksia ennakkoon.

Kaikkiaan neljää henkilöä haastateltiin, joista yksi oli Espoon kaupungin organisaatiosta, ja muut olivat haastatteluihin valittujen kaupunkien edustajia. Valitut kaupungit olivat Kuopio, Tampere sekä Jyväskylä ja valintojen perusteena painotettiin väkilukua sekä, sitä että ne sijaitsevat eri puolella Suomea. Edustajien valinnassa taas painotettiin kokemusta ja keskeisyyttä koulukuljetusten hallinnoinnista sekä siihen liittyvän datan käsittelystä. Espoon haastateltava oli Logistiikka-asiantuntija, Kuopion sekä Jyväskylän haastateltavat olivat Kuljetussuunnittelijoita, ja Tampereen haastateltava oli ryhmäkuljetusten Kuljetusesihenkilö. Haastattelujen ajankohdat ja kestot on ilmoitettu taulukossa 4.

Haastattelut toteutettiin Teams -ohjelmistolla, mikä mahdollisti niiden tallentamisen sekä litteroimisen haastateltavien suostumuksella. Ensimmäinen haastattelu suoritettiin omassa organisaatiossa, jotta voitiin arvioida haastattelun kestoa ja kysymysrungon sopivuutta sekä mahdollisia tarvittavia muutoksia tai lisäkysymyksiä. Haastattelu runko osoittautui sopivaksi, joten muutoksia ei tehty.

Kun kaikki haastattelut olivat tehty sekä nauhoitteet ja litterointi saatu niin voitiin aloittaa aineiston analysointi. Haastattelut kuunneltiin kysymys kysymykseltä läpi ja tehtiin litterointiin korjaukset. Analysoinnin apuvälineenä käytettiin Excel -ohjelmistoa, minne kirjattiin litteroinnin korjatut versiot vastauksista.

Taulukossa 7 on esitetty haastatteluihin valitut kaupungit, haastateltavien tittelit sekä haastattelujen toteutuspäivämäärät ja niiden kestot. Haastatteluaineoista kertyi yhteensä 237 minuuttia ja haastattelun keskimääräinen kesto oli 59 minuuttia.

Taulukko 7 Haastattelujen päivämäärät ja kestot

Haastattelu	**Kaupunki**	**Haastateltava**	**Päivämäärä**	**Kesto**
Haastattelu1	Espoo	Logistiikka-asiantuntuja	13.10.2023	36 min
Haastattelu2	Kuopio	Kuljetussuunnittelija	18.10.2023	43 min
Haastattelu3	Tampere (Tuomi Logistiikka)	Kuljetusesihenkilö (Ryhmäkuljetukset)	3.11.2023	89 min
Haastattelu4	Jyväskylä	Kuljetussuunnittelija	10.11.2023	69 min

3.3.2 Haastattelujen tulokset

Tässä alaluvussa käsitellään haastatteluista tehdyt havainnot sekä niistä johdetut päätelmät. Tulokset käsitellään teemoittain samassa järjestyksessä kuin haastattelutilanteissakin. Haastattelujen teemat olivat "*1. Johdanto ja koulukuljetuksiin tutustuminen*", "*2. Datan hyödyntäminen, analytiikka ja raportointi organisaatiossa*", "*3. Kehittäminen ja tulevaisuuden näkymät" sekä loppuun "Lisäykset*". Haastattelurunko ja sen kysymykset ovat nähtävissä liitteessä 1.

Teema1: Johdanto ja koulukuljetusten järjestäminen

Ensimmäisen teeman kysymyksillä 1–6 pyrittiin selvittämään haastateltavien roolia ja vastuuta koulukuljetusten järjestämisessä sekä heidän kokemustaan ja asiantuntemustaan alalta ja datan käytöstä. Lisäksi näiden kysymysten avulla haluttiin avata koulukuljetusten käytännön toteutusta, prosesseja ja mittakaavaa heidän edustamissaan organisaatioissa. Kysymysten oli tarkoitus antaa kattava kuva siitä, kuinka koulukuljetukset on järjestetty ja hallinnoitu kunkin organisaation osalta, jotta voitiin havaita yhtäläisyyksiä ja eroavaisuuksia niiden välillä.

Vaikka haastateltavat olivat kaikki keskeisiä henkilöitä koulukuljetusten järjestämisen kannalta, niin haastatteluiden vastausten perusteella voitiin havaita, että he olivat taustoiltaan ja rooleiltaan jokseenkin erilaisia. Haastateltavilla oli myös eritasoinen kokemus datan hyödyntämisestä, toisilla kattavampi ja toisilla niukempi. Kuitenkin kaikista haastatteluista huomattiin, että teknologian ja datan käyttö on keskeinen osa nykyaikaista koulukuljetusalaa, ja työntekijöiden kyky sopeutua niihin liittyviin muutoksiin on tärkeää. Lisäksi haastateltavia yhdistävä tekijä oli se, että kaikkien roolien keskiössä on kuljetusten hallinnointi sekä tiivis yhteistyö kuljetuksien kannalta keskeisten sidosryhmien kanssa.

Espoon kaupungin haastateltavana oli Logistiikka-asiantuntija, kenellä on tehtävistä parin vuoden kokemus. Sitä ennen hän on työskennellyt sekä IT-alalla että vammaistyössä, joissa molemmissa hän on hyödyntänyt dataa palvelujen kehittämisen ja laadun varmistamisen tukena. Haastateltavista pisin työkokemus koulukuljetusten parista, noin kolmenkymmenen vuoden ajalta, on Kuopion kaupungin haastateltavalla. Hän toimii roolinimikkeellä Kuljetussuunnittelija, jonka vastuu koulukuljetuksissa on laaja-alainen sisältäen muun muassa

työtehtäviä kuten hankintojen valmistelu, sopimusten kilpailuttaminen, kuljetusperiaatteiden valmistelu, kuljetuspäätösten valmistelu ja tiedottaminen sidosryhmille, taloudellinen seuranta, tietojen ylläpito sekä kuljetusten turvallisuuden varmistaminen. Tampereen osalta haastateltiin Tuomi Logistiikan ryhmäkuljetusten Kuljetusesihenkilöä. Asetelma on hieman eri muihin haasteltaviin nähden sillä Tuomi Logistiikka ei ole kaupungin yksikkö vaan logistiikkayhtiö. Se toteuttaa Pirkanmaan henkilökuljetuspalveluita mukaan lukien hyvinvointialueen kuljetukset sekä koulu- ja varhaiskasvatuksen kuljetukset. Taustaltaan Tuomi Logistiikka on kuitenkin Tampereen kaupungin vanha logistiikkayksikkö, joka muutettiin niin kutsutuksi Inhouse-yhtiöksi, missä se myy palveluita omistajilleen. Tampereen kaupunki oli Tuomen suurin omistaja hyvinvointialueiden aloitukseen asti. Tuomella haastateltava on toiminut 4–5 vuotta, joista Kuljetusesihenkilönä kahden vuoden ajan ja sitä ennen Kuljetussuunnittelijana reilu pari vuotta. Koulutustaustaltaan hän on kauppatieteiden maisteri sekä aikaisemmin työskennellyt teollisuuden parissa, missä datan käyttö on ollut myös keskeisesti mukana. Jyväskylän haastateltava oli kaupungin kuljetussuunnittelija, joka hallinnoi koulutaksiliikenteen. Rooli on vastaava kuin mitä Kuopionkin haastateltavalla, mutta työkokemus koulukuljetustehtävistä ei ole yhtä pitkä vaan reilun vuoden mittainen. Tosin kuljetusalalta hänellä on kyllä kymmenen vuoden kokemus. Datan käyttökokemus ei vielä ole hänellä suuri, mutta se on lisääntynyt koko ajan viimeisen reilun vuoden aikana, sillä dataa hyödynnetään kasvavassa määrin koulutaksiliikenteen hoitamisessa.

Espoossa koulukuljetusten kokonaisuus on jaettu kahdelle toimialalle. Kasvun ja oppimisen toimiala hallinnoi hakemuksien käsittelyn sekä tekee kuljetuspäätökset ja Logistiikkayksikkö vastaa kuljetusten toteuttamisesta. Käytännössä kuljetukset joko tilataan logistiikasta käsin tai kouluja ohjeistetaan tekemään tilaukset, joiden perusteella kuljetusyhtiöt tekevät reititykset. Kuopiossa sen sijaan koulukuljetuspäätökset tehdään kuljetussuunnittelijan toimesta samassa yksikössä mistä kuljetuksia koordinoidaankin. Reitityksestä vastaa kuljetusyhtiöt. Tampereella Tuomi Logistiikka vastaa koulukuljetusten suunnittelusta, tilausten vastaanotosta, seurannasta ja raportoinnista. He eivät tee kuljetuspäätöksiä vaan Tampereen kaupunki tekee ne, joiden perusteella koulut tilaavat kuljetukset Tuomi Logistiikalta. Jyväskylässä koulut antavat raja-arvot ja muut kuljetuksissa huomioitavat lisätiedot, joiden perusteella kaupungin kuljetussuunnittelija tekee reititykset kuljetusyhtiöille, jotka määrittävät ajoneuvokaluston kullekin reitille. Jyväskylän kaupungin toiminta eroaa reitittämisen

suhteen Espoosta ja Kuopiosta siinä, että se ei ole ulkoistanut reititystä kuljetusyhtiöille vaan hoitavat sen itse.

Kaupunkien koulutaksikuljetusten alueelliset jaot ja toimijoiden määrät vaihtelevat merkittävästi. Espoossa ominaispiirteenä on, ettei siellä ole yhtä keskusta-aluetta vaan useampi asuinkeskittymä. Esimerkiksi Etelä-Espoossa on tiivis asutus ja kouluverkko, kun taas Pohjois-Espoossa asukkaita on huomattavasti vähemmän ja koulujakin on siten vähemmän. Haastateltavan kuvauksen mukaisesti: ”*Espoo on koulupiirien mukaan jaettu neljään osa-alueeseen ja niitä sitten hoitaa neljä eri kuljetusyhtiötä.*” Kuopiossa aluejaot perustuvat oppilaaksiottoalueisiin, joita on haastateltavan mukaan yhtä paljon kuin koulujakin. Tampereella ei ole määritelty tiukkoja aluejakoja kuljetuksia toteuttavan tahon eli Tuomi Logistiikan osalta vaan heidän näkökulmassaan korostuu kuljetusten tehostaminen ketjuttamalla erilaisia kuljetuksia ja kulkemalla laajasti eri alueille. Jyväskylässä kuljetukset on jaettu alueellisesti, ja niitä hoitaa neljä kuljetusyhtiötä kaupungin koulukuljetuksissa sekä kuusi yhtiötä valtion erityiskoulun kuljetuksissa. Alla taulukossa on kuvattu haasteltavien roolit, kokemus, osuus koulukuljetusten toteutuksessa, toiminnan laajuus sekä maininta reitityksestä vastaavasta tahosta.

Taulukko 8 Haastateltavat ja vastuut

Espoo Haastateltava: Logistiikka-asiantuntija Kokemus koulukuljetuksista: n. 2 v Vastuut: Vastaa kuljetusten toteutumisesta, Tekee kuljetustilaukset tiimin kanssa tai ohjeistaa koulut tilaamaan, Kilpailuttaa kuljetussopimukset Kuljetettavia oppilaita: n. 800	**Tampere:** Tuomi Logistiikka Oy Haastateltava: Kuljetusesihenkilö Kokemus koulukuljetuksista: 4–5 v Vastuut: Suunnittelee koulujen tilaamat kuljetukset, Kilpailuttaa kuljetussopimukset Kuljetettavia oppilaita: n. 600
Kuopio Haastateltava: Kuljetussuunnittelija Kokemus koulukuljetuksista: n. 30 v Vastuut: Tekee kuljetuspäätökset, Järjestää sidosryhmien kanssa kuljetukset, Kilpailuttaa kuljetussopimukset Kuljetettavia oppilaita: n. 2200	**Jyväskylä** Haastateltava: Kuljetussuunnittelija Kokemus koulukuljetuksista: reilu 1 v Vastuut: Tekee reittisuunnitelmat kuljetusyhtiöille koulujen määrittämien raja-arvojen mukaisesti Kuljetettavia oppilaita: n. 900

Teema2: Datan hyödyntäminen, analytiikka ja raportointi organisaatiossa

Haastattelun toinen teema (kysymykset 7–18) on kattavuudeltaan ja aihealueeltaan niin sanotusti pääteema, missä tarkoituksena oli syventää ymmärrystä siitä, miten koulukuljetusten dataa nykyisellään hyödynnetään ja hallinnoidaan haastatelluissa organisaatioissa. Kysymysten avulla pyrittiin identifioimaan parhaita käytäntöjä sekä mahdollisia kehityskohteita datan hallinnassa ja käytössä.

Teeman alkuun kysyttiin, millaista dataa organisaatioissa kerätään koulukuljetuksista ja kysymykseen saatiin erilaisia näkökulmia. Espoon haastateltavan vastauksessa korostuu perusdatan, kuten matkustaja- ja matkamäärien sekä kustannusten keruu. Hän tuo kuitenkin esiin tavoitteen, että tulevaisuudessa useiden projektien myötä analysoitaisiin kuljetuksista kerättävää dataa entistä laajemmin ja tarkemmin, ottamalla jo kerättyjen peruslukumäärien lisäksi tutkinnan kohteeksi myös muita muuttujia. Erikseen hän mainitsee laadunvalvontaan liittyvän reklamaatio- ja sanktiodatan käsittelyprosessin kehittämisen, joka on nykyisellään aika manuaalinen prosessi. Tosin hän arvioi, että Espoossa ollaan laadunvalvonnan osalta kehittyneempiä moneen muuhun kuntaan verrattuna. Kuopion haastattelun vastauksen keskiössä ovat käytössä olevien paikannusjärjestelmien tarjoama data, josta saadaan tarvittavat tiedot haettua tarpeen ilmetessä. Tuomi Logistiikan haastattelussa painottuu aikataulujen noudattaminen, kustannustiedot, matkatyypit ja hiilidioksidipäästöjen seuranta. Ehkä tärkeimpänä mittarina haastateltava pitää kuljetuksen aikataulussa pysymistä niillä raja-arvoilla mitä tilaajan kanssa on sovittu. Kyseinen odotettu palvelutaso toimii marginaalina, ja yksikössä seurataan miten hyvin kuljetukset pysyvät siinä. Jyväskylän haastattelussa korostuu kustannusten, oppilasmäärien, myöhässä olevien ajojen, reittien kestojen ja oppilaiden peruutusten seuranta sekä järjestelmän hyödyntämisen ja tehokkuuden parantamisen tärkeys. Heillä on ollut käytössä reilun vuoden uusi kuljetustenhallintajärjestelmä mitä on kehitetty hyvin paljon Jyväskylän tarpeisiin. Kaikki neljä haastattelua tuovat esiin perustiedon keruun merkityksen, mutta eroavat painotuksissaan laadunvalvonnan, tarkemman datan analysoinnin ja tehokkuuden kehittämisen osalta.

Haastatteluissa ilmenee kuljetuksiin liittyvässä tiedonkeruussa ja -käsittelyssä eroavaisuuksia ja yhtäläisyyksiä. Yhteistä on pyrkimys tehokkaaseen tiedonkeruuseen ja kaikki

mainitsevatkin jonkinlaisen järjestelmän tai työkalun tiedonkeruuseen ja hallintaan. Vaihtelevuutta on käytössä olevissa järjestelmissä ja menetelmissä. Alla olevassa taulukossa on esitelty kunkin kaupungin haastatteluista ilmenneet tietolähteet, laadunvalvonta sekä erityispiirteet tiedonkeruuseen ja käsittelyyn liittyen.

Taulukko 9 Tietolähteet, laadunvalvonta sekä erityispiirteet

Kaupunki	**Tietolähteet**	**Laadunvalvonta/reklamaatiota**	**Erityishuomioita**
Espoo	Kuljetusyhtiöiden järjestelmät, Espoon kuljetustenhallintajärjestelmä sekä oppilastietojärjestelmä	Manuaalinen palautteiden tallentaminen	Dataa on kyllä tarjolla, mutta analysointi ja käsittelytyökalut on vielä kehityksessä.
Kuopio	Kihon ja Maponin -paikannusjärjestelmät sekä oppilastietojärjestelmä	Sähköposti ja Teams	Koululiitu -ohjelma koulumatkojen ja reittien vaarallisuuden arviointiin.
Tampere (Tuomi Logistiikka)	Movit -kuljetustenhallintajärjestelmä, Tuomen oma alusta	Ei mainintaa	Kuljetustensuunnittelu tapahtuu Movitisssa, mutta Tuomen oman alusta kautta voidaan välittää ruuhkapiikkien ylijäämäkuljetukset ei vakinaisesti reittejä ajaville kuljettajille. Movitista erillinen raportointipuoli minne voidaan määritellä automatisoituja raportteja.
Jyväskylä	Kuljetustenhallintajärjestelmä	Viestit kouluille, puhelut kuljetussuunnittelijalle. Myös erillinen palautekaavake.	Harkitaan palautekanavan lisäämistä sovellukseen, mikä on käytössä huoltajilla.

Dataan ja sen hyödyntämiseen liittyvistä haasteista kysyttäessä, haastateltavat tunnistavat erilaisia asioita. Espoon haastattelussa korostuu työkalujen puute ja tiedonsiirron haasteet. Haastateltavan mukaan dataa on paljon tarjolla, mutta vähemmän työkaluja millä niitä olisi kätevä analysoida. Lisäksi hän mainitsee haasteena sen, että tieto ei siirry automaattisesti järjestelmien välillä siinä määrin kuin olisi tarvetta. Kuopiossa, dataan liittyvissä ongelmatilanteissa otetaan yhteyttä suoraan järjestelmätoimittajille, jotka sitten korjaavat asian. Esimerkiksi heillä oli sellainen haaste, että paikannusjärjestelmän kahden eri tahon

samanaikaisen käytön aikana kaikkia autoja ei näkynyt, jolloin järjestelmätoimittaja joutui tekemään korjaavia toimenpiteitä. Tuomi Logistiikan haastattelun mukaan suurimmat haasteet liittyvät inhimillisiin virheisin ja käyttäjien ymmärrykseen järjestelmän käytöstä sekä järjestelmien ylläpitoon ja niiden väliseen tiedonsiirtoon. Käyttäjien tekemät merkintävirheet vaikuttavat suoraan datan laatuun ja sen hyödyntämiseen. Onkin olennaista, että auton päässä ymmärretään heidän tekemien kuittausten merkitys. Osa kuljetuksiin liittyvistä järjestelmistä ovat linkitetty “keskustelemaan” keskenään ja niiden yhteydessä tiedostetusti esiintyy ajoittain puutteita ja ongelmia. Järjestelmien ylläpito ja niiden välinen tiedonsiirto ovat jatkuvan kehittämisen kohteena. Kuten haastateltava mainitseekin: “*kun tämmöisten järjestelmien kanssa pelataan, niin se on semmoista jatkuvaa kehittämistä mitä tehdään*”. Vastaavasti miten Tuomi korostaa ajoneuvon päässä inhimillistä virhettä niin Jyväskylän kuljetussuunnittelija mainitsee, että dataan tulee jonkin verran merkintävirheitä kuljettajilta. Haastateltava mainitsee asian seuraavasta: "*esim. kuljettaja ei sulje ajolistaa niin sitten se jää niin sanotusti auki ja sitten se näyttää, että se ajo on myöhässä, vaikka se todellisuudessa ei ole myöhässä*”. Datan hyödyntämisen haasteisiin liittyen haastateltava mainitsee myös, että työajan puute on estänyt datan perusteellisen hyödyntämisen. Työtehtäviä koulukuljetusten järjestämiseen on monia ja perustoimintojen pyörittäminen, että saadaan koulukuljetukset onnistuneeksi järjestettyä, priorisoidaan kuitenkin datan hyödyntämisen edelle. Lisäksi työtehtävä haastateltavalle sekä kuljetustenhallintajärjestelmä ovat uudet, joten aikaa on mennyt paljon opetteluun. Haastateltava on oikeastaan nyt viime kevään ja syksyn aikana saanut datan enemmän hyödynnettäväksi. Haastattelujen tuloksista ilmenee erilaisia näkökulmia datan ja sen hyödyntämisen kanssa. Jokainen haastattelu valottaa eri näkökulmasta, kuinka datan hyödyntämiseen liittyvät haasteet ovat monitahoisia ja vaativat erilaisia ratkaisuja ja lähestymistapoja.

Kaikissa neljässä haastattelussa Excel nousee keskeiseksi työkaluksi, kun haastateltavilta kysytään miten ja millä ohjelmistoilla dataa analysoidaan, jotta siitä saadaan jalostettua tietoa. Lisäksi Espoon ja Tuomi Logistiikan haastatteluissa mainitaan Power BI:n käyttö, mikä viittaa datan jatkojalostukseen ja syvällisempään analyysiin. Espoon haastateltava kuvaa datan analysointiprosessia nykyistä datan käsittelyprosessia siten, että he saavat kuljetustenhallintajärjestelmän raportointityökalun avulla ladattua kuljetusdatan, jonka jälkeen se viedään jatkojalostukseen ensin Exceliin ja sitten Power BI:hin. Kuopion haastattelussa

korostuu, että tarvittavat tiedot ja toiminnallisuudet näkee jo pitkälti käytössä olevista paikannusjärjestelmistä, mutta tarvittaessa datat voi hakea niistä Exceliin käsiteltäväksi. Tuomi Logistiikan haastateltava mainitsee, että Excelin avulla hän pystyy hyvin helposti ilmentämään datasta tietoa, mutta heillä on lisäksi parhaillaan meneillään projekti missä rakennetaan Power BI ratkaisua mihin saadaan valmiiksi tietoa esiteltävään muotoon. Hän on pystynyt perustietoja käsittelemään hyvin Excelin avulla, mutta kun Power BI -ratkaisu valmistuu niin samat tiedot tulee sinne todennäköisesti suurelta osin ja siten koko raportointiprosessia saadaan vietyä automatisoidumpaan suuntaan. Lisäksi hän mainitsee, että Excelissä koordinaattipisteillä ei olla tehty mitään, mutta Power BI:ssä niistä pystytään tekemään juttuja. Jyväskylän haastattelussa ilmenee, että he eivät ole ainakaan vielä suuremmin lähteneet jatkojalostamaan dataa vaan he saavat kuljetustenhallintajärjestelmän kautta vietyä eri graafeja ja tilastoja suoraan Exceliin, mistä he katsovat tarvittavat tiedot. Analysointi tapahtuu lähinnä siten, että he lataavat vastaavan datan vuoden takaa, johon peilaten he voivat tehdä vertailua. Haastatteluja yhtenäistäväksi tekijäksi voidaan todeta, että kussakin hyödynnetään datan käsittelyssä toistaiseksi Exceliä. Eroavaisuutena on jatkojalostamisen ja automatisoinnin asteet sekä työkalujen monipuolisuus.

Haastateltavilta kysyttiin, että kuinka kehittyneenä he pitävät toteuttamaansa analytiikkaa ja samalla heille näytettiin demonstraation vuoksi kuvaa analytiikan tasojen teoriasta, mikä on esitelty tässä työssä kappaleessa 2.2.1. Espoon haastateltavan mukaan heidän tasonsa on pääasiassa kuvailevaa analytiikkaa, missä kokonaistilannetta seurataan tunnuslukujen avulla ja niitä historiadataan vertaamalla. Hän mainitsee, että ennustava analytiikka on selkeästi se kehityskohde Espoossa vielä, mitä voitaisiin hyödyntää esimerkiksi kilpailutuksissa. Suunnitteluun ja ennakointiin soveltuvia työkaluja kehitellään parhaillaan organisaation tietovarastoprojekteissa. Kuopion haastateltava ei korosta analytiikan tarvetta, mutta toteaa kuitenkin, että he tekevät ennusteita ja arvioita tulevalle lukuvuodelle muutosten perusteella. Tuomi Logistiikan haastateltava arvioi, että he sijoittuvat ennustavan ja ohjailevan analytiikan välismaastoon. He seuraavat analytiikallaan nykytilaa ja ennustavat tulevaa, jolloin analytiikka auttaa optimoimaan toimintaa ja ennakoimaan tulevia tarpeita. Optimointi ja ohjaileva analytiikka ilmenee esimerkiksi siten he tekevät laskelmia minkä avulla voivat verrata eri maksuperusteisia kuljetuksia ja valita eri aikamääreiden puitteissa optisimmat vaihtoehdot ja arvioimaan painatuksia, että jos jotakin priorisoidaan niin mitkä ovat sen seuraukset

toisaalla. Jyväskylän haastateltava arvioi, että tilastollinen analyysi, eli kuvailevasta analytiikasta kun on siirrytty ennustavan analytiikan puolelle, on se mikä osuu heidän organisaationsa tasoksi. Perusteluna hän mainitsee, että he pystyvät peilaamaan esimerkiksi tulevaan lukuvuoteen ja tulevien vuosien toimintaan saatavilla olevasta datasta ja luvuista, joten se tavallaan osuu ennustavan analytiikan esivaiheelle. Haastatteluissa löytyy yhtäläisyyksiä ennustavan analytiikan käytössä ja kuvailevan analytiikan perustassa, mikä osoittaa analytiikan tärkeän roolin osana organisaatioiden päätöksentekoa ja toiminnan suunnittelua. Eroavaisuuksia nähdään analytiikan kehittyneisyydessä ja sen käytön laajuudessa.

Kysymyksillä 12–14 oli tarkoitus selvittää, millainen raportointikokonaisuus organisaatioilla on tällä hetkellä ja millaisin mittarein, visualisoinen sekä kenen toimesta sen seuranta tapahtuu. Yleisesti ottaen kaikissa haastatteluissa korostuu kustannusten seuranta, mutta jokseenkin myös automatisaation ja prosessien tehostamisen merkitys, varsinkin Espoon ja Tampereen haastatteluissa. Espoon haastateltava mainitsee, että tällä hetkellä käytetään perusraportteja missä matka- ja matkustajamääriin sekä kustannuksiin ja palautteisiin liittyviä lukuja seurataan peruskaavioiden avulla. Tunnusluvut viedään valmiisiin taulukoihin, josta ne päivitetään Power BI:hin. Lukuja seuraa Logistiikkayksikön johto ja asiantuntijat, mutta myös koordinaattorit hyödyntävät niitä käytännön työssä kuten laskutuksen tarkastamisessa. Tulevaisuudessa raportointia on tarkoitus parantaa siten, että seurataan nykyistä useampia muuttujia kuten ajoneuvokalustoon, päästöihin sekä matkan kestoon liittyviä tietoja, ja analysoidaan niitä syvemmälle esimerkiksi toimipaikkatasolle. Tarkoitus on myös, että raporttiin saataisiin asetettua erilaisia raja-arvoja. Dataan halutaan syventyä nykyistä tarkemmin, jotta niistä voitaisiin viedä analyysejä tiedoksi eteenpäin sidosryhmille. Visualisointeina käytetään perinteisiä diagrammeja mitä Excelissä ja Power BI:ssä on tarjolla, mutta näiden lisäksi haluttaisiin ottaa mukaan monipuolisemmin visualisointeja kuten karttoja, joilla saataisiin tieto näkyväksi toiminta-alueittain muun muassa erinäisiin suunnittelu- ja ennakointitarkoituksiin. Kuopion haastateltava mainitsee heillä käytössä olevat talousraportit, mistä he saavat tarvittavat kustannustiedot eriteltynä esimerkiksi valtion selvityksiä ja vuosikustannuslaskemia varten. Tyypillisiä seurattavia tietoja ovat kuljetuskustannukset, oppilasmäärät ja eri kuljetusmuotojen määrät, mutta on mahdollista saada lisättyä muitakin tietoja mukaan seurantaan. Joukkoliikennepäällikkö ja kuljetussuunnittelija tarvitsee näitä lukuja ja ne otetaan tarkasteluun puolivuosittain syys- ja kevätlukukausi. Lopulta tarvittaessa ne

yhdistetään kokonaisen lukuvuoden tarkasteluksi. Varsinaista viikoittaista tai kuukausittaista lukujen seurantaa ei ole, vaikka he toki tarvittaessa näkevät kuukausittaiset luvut, esimerkiksi paljonko on kuljetuksia ja mitkä ovat niiden kustannukset. Haastateltava mainitsee, että ainahan visualisointeja voisi parantaa, mutta he ovat pärjänneet toistaiseksi nykyisillä tiedoilla ja järjestelmillä, missä voivat hyödyntää myös karttavisualisointeja muun muassa oppilaiden sijoittelun esittämiseen. Tuomi Logistiikan vastauksessa korostuu Power BI-raportointiprojektin valmistumisen merkitys, minkä myötä raportointi parantuu, aikaa säästyy raporttien valmistelusta ja virheen mahdollisuudet vähenevät, kun raportointiprosessia saadaan automatisoitua. Nykyisellään heillä on aikaa vievää manuaalisia vaiheita jossain määrin, koska osassa seurattavissa mittareissa joudutaan yhdistämään toisen järjestelmän tietoja. Sitten kun käytössä olevien järjestelmien data saadaan integroitua valmisteilla olevaan Power BI-ratkaisuun niin ne saadaan automatisoitua yhdeksi kokonaisuudeksi ja manuaaliset vaiheet vähenevät. Tuomilogistiikassa seurattavina tärkeinä mittareina ovat onnistumisluvut, myöhästymiset, vakiautojen käyttö ja päästöluvut. Henkilökuljetusten johto sekä tiimi käsittelevät ja seuraavat näitä tuloksia mittarista riippuen viikoittain, kuukausittain tai kvartaaleittain. Kerran viikossa haastateltava seuraa ajojen onnistumislukua, jonka avulla katsotaan kuinka paljon kuljetuksista osuu määriteltyyn marginaaliin ja kuinka paljon siitä on yli myöhästymisrajan. Sen avulla he pysyvät hyvin perillä myös mahdollisista syistä mitkä voivat aiheuttaa palautetta asiakkaiden suunnalta, jos esimerkiksi jotain radikaalia on tapahtunut kuljetuksiin liittyen tiettynä viikkona. Lisäksi hän pystyy onnistumisluvun avulla poimimaan datasta inhimilliset virheet kuten väärin tehdyt merkinnät kuljettajan päässä. Kuukausittaisessa palaverissa seurataan henkilökuljetustiimin kesken kustannuksia ja käydään läpi, miten missäkin on onnistuttu. Myös johtoryhmä käsittelee erikseen missä näiden asioiden kanssa mennään. Päästölukuja, eli paljonko on ajettu vähäpäästöisillä ja täyssähköautoilla, seurataan vielä toistaiseksi kvartaaleittain, mutta Power BI-ratkaisun myötä myös kuukausitasolla. Kerran vuodessa tehdään kokonaishiilidioksidivaikutuksen seuranta. Visualisointien paraneminen on haastateltavalle sellainen osio mitä hän odottaa Power BI-ratkaisulta. Nykyisellään he näkevät missä ajoneuvot liikkuvat milloinkin, sillä kaikki kuljetukset menevät puhelinten välityksellä myös GPS-tietoina, mutta Power BI:stä hän odottaa, että he pystyvät muun muassa heatmap-visualisoinnin avulla viemään onnistumislukuja kartalle ja pystyisivät paremmin havainnoimaan mahdolliset kaupungin alueet, minne myöhästymiset keskittyvät. Jyväskylän haastateltavan mukaan seurattavia pääasioita ovat kuljetuksen täyttöaste, kuukausittaiset kustannukset niin koko kaupungin osalta kuin kustannuspaikoittain sekä

päivittäiset kuljetusmäärät. Noin vuosi sitten käyttöön otettu uusi kuljetushallintajärjestelmä on paremmin mahdollistanut datan keräämisen ja hyödyntämisen. Koska vastaavaa historiadataa ei ole vielä pitkältä ajalta niin hyöty tulee kasvamaan lähivuosien aikana, kun dataa saadaan kerättyä. Nyt sen avulla pystytään katsomaan hyvin eri alueiden kuljetettavien määrät ja voidaan paremmin suunnitella kuljetusten yhdistelyjä sekä myös tarvittaessa tehdä ehdotuksia koulujen suuntaan aikatauluporrastuksista, jotta kuljetuskalustoa voitaisiin optimoida paremmin reiteille. Uusi järjestelmä mahdollistaa ennustettavuuden ja läpinäkyvyyden parantumisen kuljetusten liittyvässä toiminnassa. Lähtökohtaisesti lukuja seuraa kuljetussuunnittelija yhdessä työparin kanssa välillä päivätasolla, mutta suuremassa mittakaavassa kuukausitasolla. Kerran kuussa laskutuksen yhteydessä on se ajankohta, milloin he pysähtyvät seuraamaan ja analysoimaan lukuja muun muassa vertailemalla edellisvuoden lukuihin. Myös koulujen rehtorit ja sivistyspalveluiden asiaan liittyvät henkilöt, keillä on käyttöoikeudet järjestelmään seuraavat lukuja ja tarvitsemiaan tilastoja. Visualisoinneista kysyttäessä haastateltava mainitsee käyttävänsä erilaisia kaavioita seuraamistaan luvuista sekä myös karttavisualisoinnin hyödyntämisen esimerkiksi suodattamalla kartalle eri koulujen oppilaat. Hän sanoo, että ainahan visualisointeja voi parantaa, mutta he ovat tyytyväisiä tietoihin mitä he nyt järjestelmästä saavat ja että ne ovat tarpeisiin riittävät. Tavoitteena on ajan puitteissa hyödyntää nykyistä dataa mahdollisimman tehokkaasti eikä välitöntä tarvetta toistaiseksi ole uusille visualisointimenetelmille. Yhteenvetona raportoinnista voidaan todeta, että jokaisella haastattelulla on omat erityispiirteensä, mutta kaikissa tunnistetaan datan hyödyntämisen ja raportoinnin parantamisen mahdollisuudet, vaikkakin painopisteet ja kehityssuunnat vaihtelevat. Kustannusten seuranta korostuu kaikilla keskeisenä mittarina, vaikkakin seurannan tiheys vaihtelee.

Seuraavaksi haastateltavilta kysyttiin miten he hyödyntävät dataa koulukuljetuksiin liittyvässä päätöksenteossa ja kaikissa haastatteluissa datan hyödyntäminen korostuu yleisesti päätöksenteossa sekä kilpailutuksissa, mutta sen käyttö ja soveltumisalueet vaihtelevat. Espoon haastateltavan mukaan datan käyttö kilpailutusten suunnittelussa on keskeistä, mutta sen hyödyntäminen itse kuljetuspäätöksiin liittyvässä päätöksenteossa on vielä kehitysvaiheessa. Dataa pyritään hyödyntämään kilpailutuksissa mahdollisimman paljon, jotta kuljetusyhtiöiden tarjouksia pystyttäisiin optimoimaan suotuisiksi ja samalla suunnittelemaan omia resursseja. Kuljetuspäätösten ja koulupaikkojen myöntämiseen liittyen dataa pitäisi

haastateltavan mielestä hyödyntää nykyistä enemmän siten, että kuljetustarpeet huomioitaisiin jo prosessin alkuvaiheessa, mutta niin ei vielä toimita. Kuopion haastateltava mainitsee, että eri järjestelmien dataa hyödynnetään esimerkiksi koulukuljetuspäätöksissä ja sopimuskilpailutuksiin liittyvän aineiston valmistelussa. Koululiitu-järjestelmästä he näkevät koulumatkan turvattomuuden, Kiho-paikannusjärjestelmästä reitit sekä Primus-järjestelmästä oppilastiedot. Kuljetusreittien ja oppilastietojen analysointi on tärkeä osa kilpailutusten valmistelua. Tuomi Logistiikan haastateltavan mukaan datan hyödyntäminen on monipuolista kattaen sekä kaluston valinnan että kilpailutuksen suunnittelun, mukaan lukien tilitys- ja laskutusmallit. Eri alueiden erityispiirteet huomioidaan datan avulla ja haastateltava kuvaa asiaa seuraavasti: *"meillä osataan hyödyntää datan kautta sitä, että minkälaisia reittejä ajetaan missäkin ajassa, jolloin me osataan suunnitella myös niitä sopimuskilpailutuksia paremmin täsmäämään siihen alueeseen, niinku mikä sen alueen luonne on"*. Hieman muista poiketen, Jyväskylän haasteltava kokee datan käytön olevan toistaiseksi vähäistä päätöksenteossa, sillä myöntämisperusteet kuljetuspäätökselle ei ole kauhean riippuvainen datasta vaan niihin vaikuttavat muun muassa matkan pituus, matkan vaarallisuus ja mahdolliset lääkärin lausunnot. Hän kuitenkin pitää datan hyödyntämistä tärkeänä osana kilpailutusprosessia, erityisesti kalustotarpeen määrittelyssä ja liikennöitsijöiden informoinnissa. Yhteenvetona, vaikka datan hyödyntäminen on keskeistä koulukuljetusten tehokkuuden ja kustannustehokkuuden parantamisessa, sen käytön intensiivisyys ja laajuus vaihtelevat organisaatioiden välillä.

Kultakin haastateltavalta kysyttiin, että onko heillä esittää esimerkkitapausta missä datan avulla saadut oivallukset olisivat johtaneet merkittävään parannukseen tai muutokseen toiminnassaan. Espoon haastateltava mainitsi, että dataa analysoimalla on löydetty muun muassa sopimusrikkeitä ja muita virheitä, joita on pystytty oikaisemaan ja korjaamaan, mutta varsinaisia suuria muutoksia prosesseihin, ei ole vielä pystytty datan avulla tekemään. Myöskään Kuopion haastateltava ei mainitse mitään selkeää esimerkkitapausta parannuksesta, mutta korostaa yleisesti sitä, että käytössä olevista järjestelmistä on helposti ladattavissa hyvin pitkälti tarvittavat tiedot. Tuomi Logistiikan haastateltava toteaa, että on vaikeaa todeta mitään yksittäistä datan avulla saatua läpimurtoa. Pikemminkin hän näkee, että datan hyödyntäminen on jatkuvaa kehittämistä ja pikkuhiljaa parempaan suuntaan etenemistä. Toisin kuin muut haastattelut, Jyväskylän haastateltava toteaa konkreettisen esimerkin datan

hyödyntämisen johtaneesta parannuksesta. Hänen mukaansa laskutuksen kehittäminen tulee ensimmäisenä mieleen, sillä siinä työmäärä on vähentynyt huomattavasti datan hyödyntämisen kautta ja laskutus on entistä läpinäkyvämpää ja reaaliaikaista. Uuden kuljetushallintajärjestelmän avulla kaikki voidaan tarkistaa, toisin kuin aikaisemmin kun laskutus pohjautui hyvin pitkälti eri liikennöitsijöiden aineistoon, mitä oli vaikeampaa tarkistaa. Nyt he näkevät lähtökohtaisesti jokaisen reitin järjestelmästä ja niihin liittyvät kuittaukset kuljetettavista. Lisäksi haastateltavan mukaan kuljetusten yhdistely on parantunut uuden järjestelmän datan myötä, koska he näkevät nyt kartalla mistä oppilaat kulkevat ja siten he pystyvät helposti yhdistämään eri koulujen oppilaita samoille reiteille, jos se on aikataulullisesti mahdollista. Vaikka kaikissa tapauksissa on saavutettu hyötyä dataa käyttämällä, niin vain Jyväskylä erottuu mainituilla konkreettisilla esimerkeillä datan käytön myötä saaduilla muutoksilla toiminnassa.

Seuraavaksi haastateltavilta kysyttiin millaista tietoa he raportoivat eri sidosryhmille ja kaikissa haastatteluissa raportointi kohdentuu pääasiassa opetustoimeen. Espoon haastateltava mainitsi, että raportointi on nykyisin yleensä vain pyynnöstä opetuksen toimialalle eli sidosryhmälle, joka tekee Espoossa koulukuljetuspäätökset. Ensisijaisesti he ovat kiinnostuneita kustannuksista, matkustajaryhmistä sekä käyttöasteista. Kuopion haastateltava mainitsee valtionosuusselvitykset mitä varten annetaan tiedot sekä kertavuosittaisen raportoinnin kasvun ja oppimisen palvelualueelle, joka kustantaa kaikki lukuvuoden palvelusopimuksen kuljetusmenot. Heillä raportointi sidosryhmille on harvinta kaikista haastatteluista. Tuomi Logistiikan haastateltavan mukaan lähtökohtaisesti kommunikoidaan opetuksen toimialan kanssa, minne raportointi on kuukausittaista. Joka toinen kuukausi he tapaavat heidän kanssaan käymään läpi sen hetkistä tilannetta ja ajankohtaisia tapahtumia sekä jos on jotain kriittistä asiaa esillä. Kuljetusyhtiöille he lähettävät kiintiömuotoiset tilitysraportit missä on reittikohtaiset datat, että miten on kuljettu. Pyynnöstä he ovat osallistuneet myös vanhempainiltoihin, mutta muuten raportointi vanhempien suuntaan ei ole heiltä vaan kouluilta. Jyväskylässä opetustoimi saa tarvittavan datan omatoimisesti haettua kuljetushallintajärjestelmästä eikä varsinaista säännöllistä raportointia ole siihen suuntaa. Jos jotain täytyy selventää, niin silloin asiaa lähdetään tutkimaan yhteistyössä ja katsotaan, onko jotain lisäaineistoa mitä voidaan raportoida. Lähtökohtaisesti he saavat aineiston kuitenkin omatoimisesti haettua.

Haastattelujen toisen teeman lopuksi kysyttiin, että käytetäänkö reaaliaikaisen tiedon jakamiseen vanhemmille tai kouluille erityisiä alustoja tai sovelluksia. Tätä kysyttiin sen takia, sillä kohde organisaatiossa on ollut harkinnassa sellaisen ottaminen käyttöön ja haluttiin selvittää, että onko mahdollisesti sellaiset käytössä olevat koettu hyödylliseksi. Espoon haastateltava mainitsee, että he ovat tutkineet eri vaihtoehtoja esimerkiksi huoltajien tiedottamiseen, mutta toistaiseksi sopivaa ratkaisua tai sovellusta ei ole löytynyt. Heidän käytössään olevassa kuljetustenhallintajärjestelmässä on myös kouluille rajoitetut käyttöoikeudet oman toimipaikkatason tietoihin, mistä he näkevät kuljetusten tilaustiedot reaaliaikaisesti melko hyvin ja voivat halutessaan tiedottaa myös vanhempia. Kuopion haastateltava mainitsee, ettei huoltajille ja kouluille ole erillistä alustaa mistä he näkisivät reaaliaikaista tietoa vaan joko suoraan liikennöitsijät tai kuljetussuunnittelija lähettävät kuljetusreittien aikataulut kouluille, mitkä jakavat ne näkyville heidän Pedanet-sivuilleen. Tuomi Logistiikan haastateltava mainitsee, että Tampereella on käytössä Kouluun.fi, mitä kautta vanhemmat näkevät kuljetusten nouto- ja paluuajat. Siellä on mahdollisuuksia siihen, että pystyttäisiin näyttämään myös auton rekisterinumeroita tai toteutuneita kellonaikoja, mutta he ovat linjanneet, ettei niitä näytetä tietosuojasyistä. Koulujen puoleen heillä on Timit-sovellus käytössä, mitä kautta koulut pystyvät tilaamaan kuljetuksia käyttäen sitä myös vastaanottolistana sekä katsomaan milloin mikäkin auto on tulossa ja ketä on sen kyydissä. Jyväskylässä vanhemmilla on oma sovellus käytössä, josta he näkevät lapsensa tiedot siten reaaliaikaisesti, että kuittaus matkan tiedoista näkyy sovelluksessa, kun lapsi on otettu kyytiin ja kun se on matkalla sekä kun se on saapunut kohteeseen. Kouluilla on käytössä sama alusta käytössä kuin mitä kuljetussuunnittelijalla, mutta eri käyttäjäoikeuksilla. Kuljetussuunnittelijalla on pääkäyttäjäoikeudet ja kouluilla on käytössä hyvin paljon suppeammat käyttöoikeustasot, ohjaajille ja koulusihteereille omansa. Kaikissa haastatteluissa on maininta siitä, että kuljetustiedot ovat nähtävillä vanhemmille, mutta menetelmissä on eroa. Espoon ja Kuopion haastatteluissa korostuu koulujen rooli kuljetustietojen jaossa, kun taas Tampereella ja Jyväskylässä on erillinen sovellus vanhemmilla käytössä.

Teema3: Kehittäminen ja tulevaisuuden näkymät

Siinä missä Teema2 keskittyi nykyhetkeen niin Teema3 keskittyi tulevaisuuden näkymiin. Kolmannen teeman (kysymykset 19–21) tarkoituksena oli saada selville haastateltavien näkemykset siitä, miten datan rooli koulukuljetusten hallinnointiin liittyen tulee muuttumaan ja tunnistavatko he jo potentiaalisia teknologioita ja trendejä mitkä voivat vaikuttaa datan käyttöön alalla. Lisäksi pyritään tunnistamaan kohteita missä niitä pystyisi hyödyntämään.

Teeman alkuun haastateltavilta kysyttiin, miten he näkevät datan roolin kehittyvän koulukuljetusten hallinnassa tulevina vuosina. Kaikki näkevät datan roolin merkityksen kasvavan, mutta vastauksissa ei kuultu runsaasti konkreettisia näkemyksiä siitä millä tavalla. Espoon haastateltava toivoo, että työkalut kehittyvät siinä määrin, että dataa pystyttäisiin hyödyntämään entistä monipuolisemmin, jonka myötä ennakointi lisääntyisi. Hän näkee, että tämä olisi ehdottomasti yksi tapa kustannussäästöihin tulevaisuudessa. Myös Kuopion haastateltava toivoo, että kehitystä tulee olemaan, on se sitten minkälaista, vaikka hänen mielestään nykyinen järjestelmä ja kuljetusten hallinnointi on toiminut hyvin sellaisenaan. Tuomi Logistiikan haastateltava uskoo, että datan merkityksen kasvu tulee olemaan pikemminkin exponentiaalista kuin lineaarista tulevina vuosina. Kun miettii, kuinka valtavasti se on kehittynyt viimeisen kymmenen vuoden aikana siirryttäessä paperilistoista helppokäyttöiseen ja isoon määrään dataan niin odotukset ovat esimerkiksi siinä, että livedatasta pystytään korostamaan eri kustannuksia aiheuttavia haastealueita. Hän myös pohtii sitä, että tulevaisuudessa kun liikenne sähköistyy, niin pystyttäisiin suunnittelemaan ajoneuvokaluston lataukset reiteille optimoidusti dataan perustuen. Jyväskylän haastateltava sanoo, että kehitys on heillä ollut jo suuri vuoden takaiseen, uuden järjestelmän myötä. Nyt kun nähdään vaivattomasti missä kuljetusoppilaat asuvat ja mihin kouluun he ovat menossa, niin se on mahdollistanut helpomman reittisuunnittelun, kuljetusten yhdistelyn ja sopivan kokoisen autokapasiteetin hyödyntämisen. Hän näkee, että dataa voisi hyödyntää yhteistyössä koulujen kanssa siten, että voitaisiin tehdä pieniäkin lukujärjestysmuutoksia, mitkä mahdollistaisivat säästöjä kuljetuksissa, kun kuljetukset pystyttäisiin järjestää tehokkaammin.

Seuraavaksi haastateltavilta kysyttiin, että tietävätkö he olemassa olevia nousevia teknologioita tai trendejä, jotka voivat vaikuttaa miten dataa käytetään alalla tai voisivat parantaa

kuljetusten tehokkuutta. Espoon haastateltava ei osannut nimetä mitään tiettyä teknologiaa, mutta mainitsee toimintaan heijastuvina trendeinä ympäristötekijöiden huomioimisen ja asetetut ilmastotavoitteet, joita on datan hyödyntämisen näkökulmasta suuri tarve päästä seuraamaan. Kuopion haastateltavalla ei ollut tähän kysymykseen näkemystä. Tuomi logistiikan haastateltava mainitsee, että he pyrkivät olemaan mahdollisimman hyvin kehityksessä mukana ja tunnistaa, että tekoäly on nyt kovasti esillä. Vaikka hän ei juuri haastateltaessa osaa nimetä tiettyä asiaa missä sitä voisi hyödyntää niin uskoo, että fokus on automatisoinnissa esimerkiksi reititykseen liittyen. Hän myös korostaa järjestelmätoimittajien roolia, että sieltäkin pystyttäisiin määrittämään mahdollisia käyttökohteita missä tekoälyä voitaisiin hyödyntää. Lisäksi hän havaitsee haasteita tekoälyssä kuten siihen liittyvän tietoturva puolen, koska on vielä hämärän peitossa, että mitä tietoa uskaltaa syöttää tekoälylle. Jyväskylän haastateltava on henkilökohtaisesti tosi avoin kaikelle uudelle teknologialle ja trendien hyödyntämiselle, mitä voisi olla esimerkiksi 5–10 vuoden päästä. Hän näkee, että Jyväskylän koulukuljetuksissa muutos aikaisempaan on ollut suuri viimeisen reilu vuoden aikana uuden kuljetustenhallintajärjestelmän myötä, mitä tukee myös erittäin hyvä palaute kouluilta, liikennöitsijöiltä ja huoltajilta. Tekoälystä he ovat pohtineet, että sen voisi ottaa jossakin määrin vähintäänkin pilottina käyttöön lähivuosina. Hän kuvaa heidän tilannettansa tekoälyyn liittyen seuraavasti: "*Meillä tuossa periaatteessa ohjelman kautta on semmoinen pieni mahdollisuus, että voitaisiin jollakin aikajänteellä sitä hyödyntää. Siihen ei ole vielä lähdetty, koska noissa kuljetuksissa on niin paljon tiettyjä raja-arvoja, mitkä täytyy ottaa huomioon. Eri poikkeuksia muun muassa hakuaikoja oppilaille ja kaikkea muuta tämmöistä. Ollaan nähty, että se on ollut tällä hetkellä ainakin sen verran haastavaa, että siihen ei ole panostuksia lähdetty käyttämään*". Hän uskoo kuitenkin, että jollakin aika jänteellä se tulee kehitystyöhön, jotta sitä pystyttäisin vähintäänkin tietyillä alueilla hyödyntämään. Ei välttämättä vuoden päästä vaan pikemminkin 3–5 vuoden päästä,

Teeman lopuksi haastateltavilta kysyttiin, että onko heillä olemassa täyttämättömiä tarpeita tai kohteita, mille olisi mahdollisuuksia tulevaisuuden dataperusteisille innovaatioille. Espoon haastateltava mainitsee, että erilaisten data-analyysien avulla voitaisiin parantaa organisaation prosesseja varsinkin toimialojen välisessä yhteistyössä. Esimerkiksi päätöksiä tekevän ja toteuttavan tahon välillä voitaisiin datan avulla perustella päätöksiä ja ohjata toimintaa oikeaan suuntaan. Siinä on ollut haastetta ajoittain ja haastateltava odottaa, että data-

analytiikkatyökalujen- ja projektien myötä toimialojen välisiä prosesseja saataisiin parannettua. Tekoälyn ja robotiikan hän mainitsee mahdollisuuksina, ja vaikka aiheeseen liittyviä projekteja on Espoon kaupungilla useita käynnissä niin ne eivät ole vielä laajasti tulleet osaksi logistiikan vastuualueisiin. Kuopion haastateltava toteaa, että he ovat tulleet toimeen nykyisillä tiedoilla ja ilmaisee huolensa lisäkustannuksista, mikäli uusia järjestelmiä tai ohjelmia lähdetään ottamaan käyttöön. Myöskään Tuomi Logistiikan haastateltava ei anna selkeää näkemystä kysymykselle vaan toteaa sen olevan vähän hämärän peitossa. Sen sijaan Jyväskylän haastateltava mainitsee dataperusteisen innovaation kohteena ennustettavuuden. Hän pohtii, että jos sitä saataisiin jollakin tavalla lisättyä, niin olisi hyvä koska se on tärkeä asia toiminnan kannalta. Konkreettisena toimena hänellä tulee mieleen oppilasmäärien muutosten ennustaminen, koska ne vaikuttavat muun muassa kilpailutusten suunnitteluun. Haastateltava näkee uudet innovaatiot mahdollisuutena ja ottaa avoimesti pienessä mittakaavassa kokeiluun pilottina, nähdäkseen miten ne voisivat heidän tehtävissään toimia. Lopuksi tätä näkemystä tukien hän mainitseekin: "*Kyllä kuitenkin työ ja alat muuttuvat, niin pitää olla valmis siihen muutokseen myös*".

Lisäykset

Haastattelun lopuksi on Lisäykset-osio (kysymykset 22 ja 23), jonka tarkoituksena on nimensä mukaisesti tarjota mahdollisuus vielä tuoda esille lisäkommentteja haastatteluun liittyvistä aiheista, mikäli jotakin olennaista ei tullut mainituksi aikaisempien teemojen kysymysten yhteydessä. Vastaajia pyydetään myös kertomaan suosituksia muille organisaatioille datan hyödyntämiseen liittyen.

Espoon haastateltava mainitsee, että hänellä ei ole muuta lisättävää datan roolista heidän toiminnassaan kuin se, että työkaluilla sekä järjestelmillä on suuri merkitys. Kun ne ovat mahdollisimman tarkoituksenmukaisia ja toimivia niin silloin aidosti kaikki data mitä on saatavilla, on hyödynnettävissä. Suosituksena muille kunnille hän nostaa sen, että kannattaa ottaa käyttöön sellainen kuljetustenhallintajärjestelmä mikä mahdollistaa, että koulut itse pystyvät hallinnoimaan omien kuljetusoppilaidensa kuljetuksiin liittyviä muutoksia, tilauksia sekä pystyvät toimittamaan aikataulut suoraan järjestelmään. Näin kuljetuksia hallinnoivan yksikön toiminta kevenee ja tällaisen järjestelyn käyttöönotosta hän positiiviseen sävyyn

kuvaakin: "*koulut ovat yllättävän nopeasti sitten loppupeleissä kuitenkin omaksunut tämän käyttöönsä, vaikka se on aika erilainen työtapa kuin mitä oli vielä vaikka 3–4 vuotta sitten*". Järjestelmästä on saatu todella paljon hyötyä esimerkiksi käytännön työhän niille koulujen ohjaajille, jotka valvovat taksiin menijöitä ja taksilla tulijoita päivittäin. Prosessista jää pois monta välikättä, kun ohjaajat voivat itse tehdä tarvittavat muutokset järjestelmään sekä katsoa ja merkitä saapuneet oppilaat. Kuopion haastateltava loppuun mainitsee, että heillä käytössä olevat paikannuslaitteet, mistä saa tarvittavaa tietoa, ovat olleet hyvät sekä heille että toimittajille. Näitä hän voikin suositella myös muille toimijoille vastaaviin tehtäviin. Yleensäkin hän suosittelee hankkimaan hyvät järjestelmät, mistä pystyy seuraamaan tarvittavaa tietoa ja saa sitä mahdollisimman jalostettuna, jolloin ei itse tarvitse työstää raporttia, mikä helposti on aikaa vievää. Tuomi Logistiikan haastateltava pohtii lopuksi, että koulukuljetusten järjestäminen on työtä mikä on myös veronmaksajien rahoilla, joten ei alalla tehdä päätöksiä tunnepohjalta vaan data on vahvasti tukena päätöksenteon taustalla. Suosituksena muille hän sanoo, ettei dataa kannata ajatella hirveänä "mörkönä" kuten moni sen semmoisena mieltää. Organisaation työntekijöiden mielikuvat ja asenteet vaikuttavat tosi paljon siihen, miten asioihin kuten datan mahdollisuuksiin suhtaudutaan. Hän kuvaa yhä asiaa seuraavasti: "t*osi paljon riippuu työntekijöiden kiinnostuksesta, että miten se data koetaan ja että onko se niin kun hyvä vai huono, pelottava muodoton massa mistä ei oikein saa kiinni*". Hän painottaa myös, että työntekijöiden suhtautuminen on tosi kriittinen siihen, osataanko dataa hyödyntää ja nähdäänkö siinä mahdollisuuksia vai koetaanko se turhana lisätyönä, joka ei johda mihinkään. Sen sijaan, että koettaisiin että ostetaan tai hankitaan dataa ja raportteja vain sen takia, että ne lähetetään pomolle, joka kuitenkin laittaa ne silppuriin niin asiat on myös mahdollista nähdä palkitsevina. Asiat voidaan kokea positiivisemmin siten, että nähdään datalla olevan suora vaikutus tulevaisuuden suuntiin. Haastateltavan suosituksen pääpointtina on, että kannustettaisiin datamyönteisempään toimintaan ja, että se nähtäisiin enemmän mahdollisuutena kuin esteenä tai turhana asiana. Jyväskylän haastateltava lopuksi mainitsee datan roolista, että se on hänen mielestään erittäin mielenkiintoinen osa-alue työnkuvassaan ja että datasta varmasti saadaan enemmän hyötyä tulevina vuosina ja sen merkitys tulee korostumaan. Hän uskoo, että datan avulla voidaan parantaa entisestään palvelun laatua ja ohjaamaan kuljetuskalustoa tehokkaammin sekä parantaa yhteistyötä liikennöitsijöiden kanssa. Paremman datan hyödyntämisen pohjalta kaikki osapuolet hyötyvät siten, että kaupunki saa kustannustehokkaampaa palvelua, liikennöitsijät ennakoivampaa toimintaa ja kilpailukykyisempiä kuljetuksia sekä koulut ja huoltajat ennakoivaa, hyvin toimivaa ja

luotettavaa palvelua. Haastateltavan mukaan datasta saadaan jo nyt tiettyjä asioita poimittua hyvin toimintaan, mutta hän uskoo, että sitä pystytään hyödyntämään vieläkin paremmin. Nykyisen lukuvuoden koulukuljetukset ovat menneet Jyväskylässä todella hyvin syksyn osalta, sillä reklamaatioita on tullut erittäin vähän ja kuljetukset ovat pysyneet suhteellisen hyvin aikataulussa. Haastateltavan mukaan palvelu on toiminut ja datan hyödyntäminen on ollut yksi osatekijä onnistumisessa. Suosituksena muille organisaatioille haastateltava toteaa, että kuljetustenhallintajärjestelmä on sellainen asia mihin kannattaa ehdottomasti panostaa. Esimerkiksi sen ansioista koulukuljetukset on saatu läpinäkyväksi ja he pystyvät katsomaan järjestelmästä muun muassa miten reitit ovat myöhässä ja milloin kuljetukset aloitetaan. Myös huoltajilla on järjestelmään kytkeytyvät sovellus mistä he näkevät kuljetuksien tiedot. Ennen kuljetustenhallintajärjestelmän käyttöönottoa toiminnassa soitettiin paljon eri sidosryhmien välillä joka suuntaan, mikä on vähentynyt järjestelmän käyttöönoton jälkeen huomattavasti, kun tieto on läpinäkyvää ja monessa tilanteessa reaaliaikaista. Haastateltavan mielestä kaupungit voisivat tehdä enemmän yhteistyötä, koska monet kaupungit varmasti kohtaavat samoja haasteita ja ongelmia. Siten organisaatiot voisivat jakaa ja löytää hyviä toimenpiteitä. Vaikka jonkin verran hänelle on jo tullut yhteydenottoja muilta kaupungeilta, niin se on ollut kuitenkin melko vähäistä. Haastateltava sanoo uskovansa, että yhteistyön parantaminen yksi tavoite mihin itse koittaa pyrkiä jollakin aikajänteellä, jotta hän löytäisi eri kaupungeista kollegoita kenen kanssa voisi vaihtaa ajatuksia ja kuunnella välillä kuulumiset, siitä miten muilla menee ja voitaisiinko toinen toiselta oppia kenties jotakin.

3.4 Osa II: Datamalli

Tämän tutkimuksen kappale 3.4 käsittää työn empiirisen osuuden toisen osan, missä rakennetaan kirjallisuuskatsaukseen, haastattelujen tuloksiin sekä kohdeorganisaation esittämiin toiveisiin pohjautuva datamalli, minkä on tarkoitus toimia työkaluna päätöksenteon tukena. Datamalli on myös osana laajempaa tietoallashanketta ja sen on tarkoitus olla Logistiikkayksikön käytössä oleva alkuvaiheen versio, minkä avulla voidaan havaita mahdollisuuksia ja haasteita mitä saatavana olevasta datasta on mahdollista laskea sekä myös verrata että tulokset täsmäävät datamallin ja kehitteillä olevan tietoallashankkeen havaintojen ja laskelmien välillä. Datamalli toteutettiin marraskuun 2023 sekä maaliskuun 2024 välisenä aikana. Sen

rakentamissa on OSEMN-prosessimallia ja alakappaleissa on esitetty mihin vaiheeseen prosessista kappale liittyy.

3.4.1 Sovelluksena Power BI

Markkinoilla on useita alustavaihtoehtoja analytiikan ja Business Intelligence:n toteuttamiselle. Kuvasta 11 nähdään arvostettu Gartnerin Magic Quadrant vuodelta 2023, missä alustapalveluiden tarjoajat ovat sijoitettuna suorituskyvyn ja vision täydellisyyden perusteella neljään kategoriaan: Johtajat, Haastajat, Visionäärit ja Oman markkinansa omaavat. Huomataan, että Microsoft on selkeä markkinajohtaja, jota lähimpänä seuraavat Salesforce (Tableau) sekä Qlik. (CXToday 2023)

Kuva 11 Gartnerin 2023 Magig Quadrant Analytiikan ja Business Intelligence -alustoista (CXToday)

Tämän työn datamallin toteuttamiseen valittiin Microsoft Power BI Desktop -sovellus, minkä lataaminen ja käyttäminen on maksutonta. Power BI Desktop on ollut Espoon kaupungilla jo useamman vuoden käytössä niin talousraportoinnissa kuin eri toimialojen omissa raportoinneissaan. Myös Logistiikkayksikössä on aiempaa käyttökokemuksesta sovelluksesta, joten valinta oli perusteltu eikä nähty syytä lähteä kokeilemaan toista sovellusvaihtoehtoa. Työn empiirisessä osassa kehitetyn datamallin kehittämisessä hyödynnettiin aiempaa käyttökokemusta ja osaamista Power BI:stä sekä verkosta saatavilla olevia ohjeita.

Power BI Desktop on sovellus, minkä avulla voidaan muodostaa yhteys datoihin sekä muuntaa ja visualisoida niitä. Muodostamalla yhteys datalähteisiin ne voidaan yhdistää datamalliin. Tätä kutsutaan usein mallintamiseksi. Datamallin avulla voidaan luoda visualisointeja ja visualisointikokoelmia, joita pystytään jakamaan raportteina muiden organisaation jäsenten kanssa. Yleisimmät Power BI Desktopin käyttötavat ovat dataan yhdistäminen, datan muuntaminen ja siistiminen datamallin luomiseksi sekä visualisointien ja raporttien luominen. Raporttien jakaminen muiden kanssa kuitenkin vaatii maksullisen lisenssin Power BI -Serviceen. (Microsoft 2023)

Raportti (Report View), Data (Data View) ja Malli (Model View) -näkymä ovat Power BI Desktopissa on kolme käytettävissä olevaa näkymää. Raporttinäkymässä voidaan luoda raportteja sekä visualisointeja, ja siinä käytetään myös eniten aikaa luomistyöstä. Datanäkymässä nähdään raporttiin liittyvässä datamallissa käytetyt taulukot, mittarit sekä muun datan ja voidaan muuntaa ne parhaalla tavalla raportin mallia varten. Mallinäkymässä voidaan nähdä datamallin taulukoiden väliset suhteet sekä hallita niitä. Datamallin luominen tapahtuu Power Query -editorin avulla, missä pystytään siistiä ja muuntaa malliin noudettua dataa. Editori tallentaa jokaisen tietojen muuntamisen vaiheen ja aina kun tämä kysely muodostaa yhteyden tietolähteeseen niin nämä vaiheet suoritetaan, jolloin data muotoillaan aina määritetyllä tavalla. (Microsoft 2023)

3.4.2 Tavoitteen määrittäminen

Työssä Power BI -sovelluksella toteutettavan datamallin tavoitteena on luoda analyyttinen työkalu mitä Espoon kaupungin Logistiikkayksikkö voi hyödyntää koulukuljetuksiin

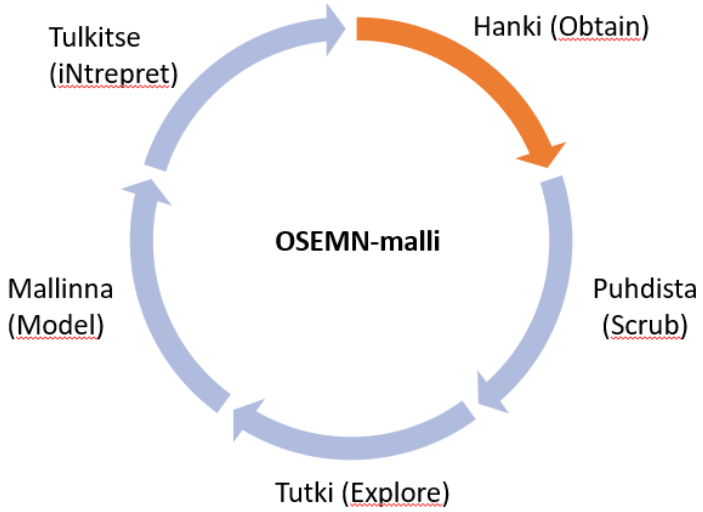
liittyvien työtehtäviensä ja päätöksenteon tukena. Ensisijaisena käyttäjänä on laajemman käyttötarpeen omaava Logistiikka-asiantuntija, mutta muita tunnistettuja mahdollisia käyttäjiä ovat Logistiikkajohtaja, Kuljetuspäällikkö, Kuljetuskoordinaattori, Palvelukoordinaattori sekä Projektikoordinaattori. Niin Ad hoc -tehtävät kuin myös laajemmat selvitystyöt sekä kilpailutusaineiston valmistelut ovat esimerkkejä käyttökohteista mihin työkalua voidaan hyödyntää. Datamallia voidaan myös käyttää hyödyksi laajemman tietoallaskehitys-hankkeen tukena esimerkiksi aineistojen ja laskelmien vertailussa ja paikkansapitävyyden tarkistamisessa.

Tähän mennessä kohdeyksikössä ongelmana on ollut päätöksenteon tukena olevien mittaristojen ja visualisointien puute. Yksikössä tarvitaan työkalua helpottamaan suunnittelua ja päätöksentekoa esimerkiksi kilpailutusten yhteydessä datavisualisointien ja mittarien avulla. Toistaiseksi lukuja on kerätty kuukausitasolla erillisiin seuranta-Exceleihin noutamalla luvut eri lähteistä kuukausittain, mutta tietojen visualisointi on ollut melko vähäistä. Lisäksi lukuja on haettu tarpeen mukaan erilaisiin projekteihin ja selvityksiin. Työssä luotavan datamallin tarkoitus on olla ratkaisu ongelmaan ja toimia kuljetusdataan perustuvana analyyttisena työkaluna, mistä nähdään visualisointien avulla datan luonne ja siten tehdä perustellumpia päätöksiä ja aikaansaada kustannussäästöä. Lopputuotteen myötä raportointi parantuu ja kuljetuslukujen seuranta helpottuu.

Lisäksi datamallin tavoitteena on edesauttaa ja olla osana Espoon logistiikan laajempaa tavoitetta, missä useiden projektien ja hankkeiden myötä pystytään hyödyntämään ja analysoimaan tarkemmin saatavilla olevaa dataa. Käytännössä se tapahtuu parantamalla visualisointeja jo entuudestaan seurattujen perinteisempien muuttujien, kuten kuljetusten ja matkustajien lukumäärät osalta sekä ottamalla tutkinnan kohteeksi lisäksi myös muita muuttujia mitä ei olla aikaisemmin seurattu. Visualisointeihin on tarkoituksen mukaista tuoda porautumismahdollisuus, jotta voidaan tarkastella dataa tarvittaessa toimipaikkatasolla asetetuin raja-arvoin. Kuljetusdata on tarkoitus tuoda karttanäkymälle, missä voidaan tehdä tarkastelua analyysin tukena ja vertailla Espoon sisäisesti eri alueita ja kouluja. Myös kuljetusten laadunvalvontaan liittyvä data, mitä kerätään manuaalisesti asiakaspalvelun kautta, otetaan osaksi datamallia täydentäväksi aineistoksi.

3.4.3 Data-aineiston hankkiminen ja kuvaus

Tässä kappaleessa kuvataan datamalliin käytetyt data-aineistot sekä niiden hankkiminen mikä on OSEMN-mallin ensimmäinen vaihe (kuva 12). Datamalli koostuu kahden datalähteen aineistosta, kuljetushallintajärjestelmän raportointityökalun kuljetusdatasta sekä asiakaspalvelun manuaalisesti keräämästä ja ylläpitämästä reklamaatioraportin datasta. Ensimmäisenä mainittu toimii pääaineistona ja jälkimmäinen lisäaineistona. Dataan tutustuminen ja niiden hankkimiseen vaaditut toimenpiteet selvitettiin työn alkuvaiheessa, jolloin pystyttiin hahmottamaan dataan liittyvät rajoitukset ja erityisominaisuudet. Tutkijan aiempi työkokemus koulukuljetuksista ja siihen liittyvistä järjestelmistä nopeutti data-aineiston hankkimisen vaihetta.



Kuva 12 Data-aineiston kuvauksen ja hankkimisen yhteys OSEMN-malliin

Pääaineistona käytettiin kuljetushallintajärjestelmän dataa, mikä saatiin ladattua kappaleessa 3.2.3 kuvatusta järjestelmätoimittajan raportointityökalusta. Data oli toimittajan SQL-kyselyn avulla toteuttama kustomoitu taulukkomuotoinen raportti, minkä Tilaaja (Espoon kaupunki) määritti tutkimuksen alkuvaiheessa ja tilasi saataville raportointityökaluun. Sieltä raportit saatiin ladattua kuljetusyhtiökohtaisesti halutulta aikaväliltä. Kunkin kuljetusyhtiön raportti on samassa muodossa ja niissä on samat sarakkeet käytössä, mikä on edellytys datamallin rakentumiselle ja toimivuudelle. Raportointityökalussa muodoksi valittiin XLSX

sijaan CSV, koska se on yleensä pienempi ja tehokkaampi suurille tietomäärille ja tyypillisesti nopeampi ladata yksinkertaisuuden vuoksi. Lisäksi raporttien datassa ei ole pudotusvalikoita, kommentteja, laskukaavoja eikä muita edistyneitä ominaisuuksia, joten ei ollut tarkoituksen mukaista ladata .XLSX-muodossa raportteja. CSV-muotoinen data ladattiin kuukausiraportteina vuoden 2023 tammikuusta vuoden 2024 maaliskuuhun asti. Raporttien aikaväliksi valittiin kuukausi, koska tarkastelua tehdään lähtökohtaisesti kuukausittain, joten on selkeämpää kerätä ja ylläpitää aineistoa kuukausitasolla. Tallennustilaksi datamallille luotiin lähdekansion alakansio, minne aineisto vietiin. Jatkossa kansioon voidaan yksinkertaisesti tallettaa tulevien kuukausien vastaavat raportit ja päivittää Power BI:hin. Raportissa sarakkeita on yhteensä 26 ja ne on esitetty taulukossa 10. Raportti sisältää myös asiakkaiden kotiosoitteet ja niiden koordinaatit, joita ei voitu tietosuojasyistä tässä työssä sen tarkemmin mainita.

Taulukko 10 Raportointityökalusta saadun datan rakenne

Nro	Sarake	Muoto	Selite
1.	**Päivämäärä**	18.1.2023	Tapahtuma-ajan päivämäärä
2.	**Vuosi**	2023	Tapahtuma-ajan vuosi
3.	**Kuukausi**	1	Tapahtuma-ajan kuukausi
4.	**Viikko**	3	Tapahtuma-ajan viikko
5.	**Matkustaja**	Sukunimi Etunimet	Matkustajan nimi
6.	**Alkuaika**	tammikuu 18, 2023, 09:10	Kuljetuksen alkamisaika
7.	**Päätösaika**	tammikuu 18, 2023, 09:25	Kuljetuksen päättymisaika
8.	**peruttu**	true/false	Merkintä kuljetuksen perumisesta
9.	**Peruttu myöhässä**	true/false	Merkintä kuljetuksen myöhässä perumisesta, jolloin kuljetusyhtiö voi veloittaa kuljetuksen
10.	**Ei saapunut**	Ei saapunut	Merkintä jos kuljetettava jätti saapumatta kuljetukseen
11.	**Apuvälineet**	Pyörätuoli	Merkintä kuljetettavan mukana olevista apuvälineistä
12.	**Saattaja**	Saattaja	Merkintä onko kuljetettavalla saattaja mukana kuljetuksessa
13.	**Kuljetuspäätös**	Koulukuljetus sisäinen	Kuljetuspäätöksen nimi
14.	**Hinta kuljetuspäätöksellä**	10.00	Kuljetettavan kuljetuspäätökseen perustuva matkan hinta
15.	**Kustannuspaikka**	//xxxxxxxx//	Tiliöintiavain, mihin kustannukset kohdentuvat
16.	**Liikennöitsijä**	Taksiyhtiön nimi	Kuljetusyhtiön nimi
17.	**Pääkohde**	2022 Karhusuon koulu	Toimipaikan nimi
18.	**Kohteen osoite**	Karhuniitynkuja 3	Toimipaikan katuosoite
19.	**Kohteen postinumero**	02810	Toimipaikan postinumero
20.	**Kohteen kunta**	Espoo	Toimipaikan kunta
21.	**Kohteen leveysaste**	60.22855095	Toimipaikan koordinaatin leveysaste
22.	**Kohteen pituusaste**	24.6210161489306	Toimipaikan koordinaatin pituusaste
23.	**Kodin osoite**	Kodinkatu 10	Kuljetettavan kotisosoite
24.	**Kodin postinumero**	02810	Kuljetettavan kodin postinumero
25.	**Kodin kunta**	Espoo	Kuljetettavan kodin kunta
26.	**Kodin leveysaste**	60.22855095	Kuljetettavan kotisosoitteen koordinaatin leveysaste
27.	**Kodin pituusaste**	24.6210161489306	Kuljetettavan kotisosoitteen koordinaatin pituusaste

Toinen datamallin lähdeaineisto on peräisin reklamaatioraportin taulusta, mikä sisältää taulukon 11 mukaiset sarakkeet. Reklamaatioraporttia hyödynnetään kuljetusten palvelunlaadun seurannassa. Tässä työssä kyseinen data toimii toissijaisena ja pääaineistoa tukevana aineistona. Kyseinen data kopiotiin asiakaspalvelun manuaalisesti keräämästä ja ylläpitämästä reklamaatioraportin datasta uuteen datamallia varten luotuun reklamaatio-Exceliin, mistä se noudettiin Power BI:hin. Data ladattiin koko vuoden 2023 osalta ja jatkossa kuukausikohtaisesti vuodelta 2024 aina kuukausilukujen seurannan yhteydessä. Reklamaatiodatan avulla voidaan seurata reklamaatioiden osuutta toteutuneisiin kuljetuksiin suhteutettuna, mikä on yksikön entuudestaan seuraama mittari. Sarake "*Reklamaation tekoaika (pvm)*" toimii yhdistävänä tekijänä kuljetushallintajärjestelmän pääaineistoon, jolloin voidaan seurata reklamaatioprosenttia kuukausitasolla. Muita yhdistäviä tekijöitä ei tämän työn puitteissa pääaineistoon luotu. Taulukkojen välillä termit ovat kirjoitusasultaan jokseenkin eriäviä eikä yhteistä tunnuskoodistoa ole.

Taulukko 11 Reklamaatioraportista saadun datan rakenne

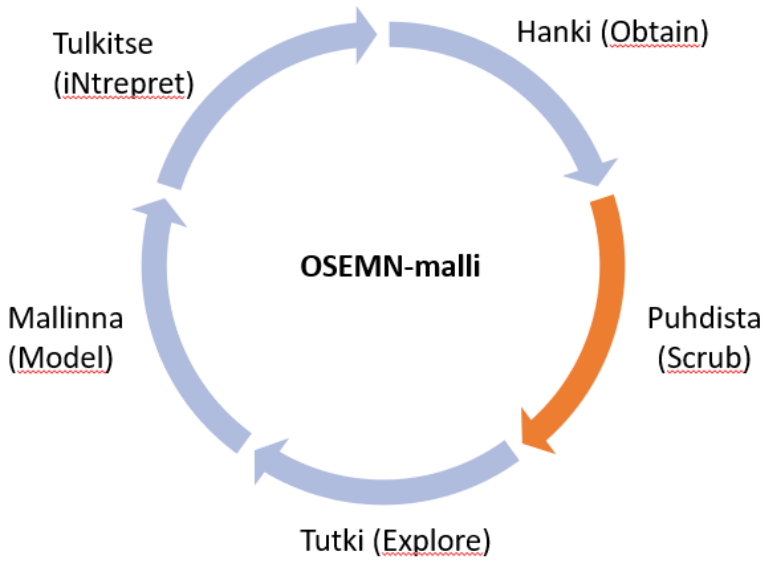
Nro	Sarake	Muoto	Selite
1.	**Reklamaation tila**	Valmis	Joko valmis tai käsittelyssä
2.	**Reklamaation tekoaika (pvm)**	10.1.2023	Päivämäärä, jolloin reklamaatio on ilmoitettu
3.	**Reklamaationumero**	1	Juokseva reklamaatiokohtainen numero
4.	**Reklamaation lähettäjä**	Etunimi Sukunimi	Reklamaation ilmoittajan nimi
5.	**Reklamaation aihe**	Myöhästyminen (15min)	Esim. myöhästyminen, asiakasta ei haettu tai kuljetusaika ylittyi (60min)
6.	**Vakavuus**	Tavanomainen	Vakava, kärjistyvä, tavanomainen, vähäinen, perusteeton, kiitospalaute
7.	**Kuljetettavien lukumäärä**	3	Montako kuljetettavaa reklamaatio koskee
8.	**Kuljetusyritys**	Kuljetusyhtiö1	Kuljetuksia hoitavat sopimustoimittajat
9.	**Kuljetuslaji**	Koulukuljetus	Koulukuljetus, Varhaiskasvatus, Kipsikallet, Lomahoito
10.	**Toimipiste**	Jupperin koulu	Toimipiste mitä reklamaatio koskee
11.	**Tapahtuma-aika**	10.1.2023	Päivämäärä, jolloin reklamaatio tapahtunut

12.	**Lisätieto**	Oppilaat 1, 2 ja 3 myöhästyivät koulusta 30minuuttia	Täsmennyksiä kuten tapahtuman paikka ja asiakas
13.	**Vastine Pyydetty (pvm)**	10.1.2023	Päivämäärä, jolloin kuljetusyhtiöltä on pyydetty vastine reklamaatiosta
14.	**Vastine Saapunut (pvm)**	15.1.2023	Päivämäärä, jolloin kuljetusyhtiöltä on saatu vastine reklamaatiolle
15.	**Onko sanktioitava**	Kyllä / Ei	Merkintä onko reklamaatio sanktioitava
16.	**Sanktiosta ilmoitettu ky:lle (pvm)**	4.2.2023	Päivämäärä, jolloin sanktio on lähetetty kuljetusyhtiölle
17.	**Ky vastasi ilmoitukseen (pvm)**	10.2.2023	Päivämäärä, jolloin kuljetusyhtiö vastasi ilmoitukseen
18	**€/hlö**	50	Sanktion suuruus per kuljetettava
19.	**KPL ajo-tapahtumia**	3	Montako kuljetettavaa sanktio koskee
20.	**Yhteensä**	150	€/hlö x KPL ajotapahtumia
21.	**Onko sanktiolasku lähetetty?**	On Helmikuussa	Maininta missä kuussa jos sanktio on lähetetty

Työssä käytettyjen datojen tallennustilana toimi Tutkijan oman tietokoneen työpöytä, koska data sisälsi henkilötietoja mitä ei voinut tietoturvasyistä jakaa Sharepoint -alustalle. Työpöydällä luotiin kansio datamallille ja alakansiot sen lähteenä käytetylle datoille, mistä ne noudettiin Power BI:hin. Raporttien yhdistäminen tapahtui Power BI:n Power Query:ssä.

3.4.4 Datan puhdistus ja muuntaminen

Tässä kappaleessa kuvataan datan puhdistamiseen ja muuntamiseen liittyvät toimenpiteet. Suurin osa puhdistamistoimenpiteistä tehtiin työn alkuvaiheessa datan noudon yhteydessä, mutta joitakin muokkauksia tehtiin myös myöhemmin tarpeen vaatiessa esimerkiksi datan tutkimisen vaiheessa, kun ensimmäiset mittarit ja visualisoinnit oli luotu. Datan puhdistus on OSEMN-mallin toinen vaihe (kuva 13).



Kuva 13 Datan puhdistuksen ja muuntamisen yhteys OSEMN-malliin

Datan esikäsittely ja puhdistustoimenpiteet tehtiin Power BI:n Query editorissa pääaineiston osalta ja reklamaatiodatan osalta sille luodussa Excel-työkirjassa. Aineistojen tallennustilana toimivaan lähdekansioon luotiin oma alakansio raportointityökalun datalle sekä reklamaatiodatalle, mitkä yhdistettiin sieltä suoraan Power BI:hin.

Kuljetustenhallintajärjestelmän raportointityökalun data

Raportointityökalun dataa tarvittiin lähes jokaisen mittarin tiedon tuottamiseen. Ainoastaan suoraan reklamaatiodatasta saatujen määrien laskemiseen sitä ei tarvittu. Lataaminen Power BI:hin ja kyseisten taulukoiden yhdistäminen tapahtui seuraavien askeleiden mukaisesti:

1. Raportointityökalusta ladataan CSV-muodossa kunkin kuljetusyhtiön kuukausikohtainen data tammikuusta 2023 alkaen ja viedään työpöydällä sijaitsevaan Datamalli -kansion lähdedatan alakansioon.

2. Power BI:ssä painetaan Noudetaan tiedot ja valitaan oikea lähde -kansio missä data sijaitsee. Esikatselusta valitaan Muunna tiedot.

3. Muut sarakkeet poistetaan paitsi Content ja Name. Content -sarake sisältää kaiken datan ja Nimi -sarake näyttää tiedoston nimen, mikä on hyvä säilyttää mahdollisten virheiden paikantamisessa.
4. Sitten luodaan Mukautettu sarake, jolle määritettiin nimi DataLoad ja sen Mukautetuksi sarakekaavaksi: Table.PromotoHeaders(CsvDocument([Content])). Tämä lataa datataulut Content -sarakkeesta yhdistämistä varten, korottaen kunkin taulukon ensimmäisen rivin tarkoituksenmukaisesti otsikoiksi. Nyt kun data on ladattuna mukautetussa sarakkeessa DataLoad niin Content -sarake voidaan poistaa ja Dataload -sarakkeen taulukot yhdistää laajentamalla sarake ja valitsemalla pois valinta "Käytä alkuperäisen sarakkeen nimeä etuliitteenä".
5. Datataulukoiden yhdistämisen askeleet on nyt tehty ja taulukoiden sarakkeet näkyvät laajennettuna, jonka jälkeen valitaan tarvittavat tietotyypit kullekin sarakkeelle. Tämän jälkeen Query editor suljetaan ja otetaan käyttöön yhdistetty data.
6. Nyt kun tietolähteenä toimivaan lähdekansioon tallennetaan seuraavan kuukauden raportit ja painetaan päivitä Power BI:ssä, niin kyseinen data päivittyy raportille.

Datan puhdistus ja muuntaminen alkoi sen lataamisen jälkeen muuntamalla sarakkeiden tietotyypit sopiviksi ja korjaamalla datasta löytyviä virheitä kuten joitakin desimaalivirheitä luvuissa. Tällaiset virheet korjattiin oikeiksi Korvaa arvot -toiminnolla mikä luo vastaavan ehdon kuten seuraavassa esimerkissä on esitetty: "*Table.ReplaceValue(#"Korvattu arvo", 1210,12.1,Replacer.ReplaceValue,{"Hinta kuljetuspäätöksellä"})*" Kaikki tunnistetut virheet korjattiin, mutta niitä palattiin korjaamaan myös myöhemmässä vaiheessa työn edetessä aina kun sellainen huomattiin. Lähtödatassa ei katsottu olevan turhia sarakkeita, joten mitään ei poistettu vaan taulukko pidettiin alkuperäisessä muodossa. Kaikkien Pääkohde -sarakkeessa sijaitsevien toimipisteiden nimien etuliitteenä oli vuosiluku "2022", mikä poistettiin korvaamalla sarakkeesta kyseinen arvo tyhjällä merkillä. Matka-aikojen kestoja varten luotiin uusi mukautettu sarake, missä sarakekaavassa lasketaan [Päätösaika] - [Alkuaika]. Koska matka-ajan muoto on "00:0:00:00.0000000" sitä haluttiin vielä tehdä helppolukuisammaksi luomalla siitä yhä toinen mukautettu sarake sarakekaavalla Duration.TotalMinutes([Matka-aika]). Tämän luodun sarakkeen avulla voitiin sitten mallinnusvaiheessa seurata kuljetusten kestoja. Kun matka-ajat olivat sarakkeessa niin sen avulla pystyttiin luomaan

myös uusi sarake “Matka-aikaluokitus” minkä avulla eri kestoiset matka-ajat voitiin segmentoida eri luokkiin. Sarakkeen arvojoukkona toimii luokituksiin asetetut aikavälit, jotka olivat “Alle 50 min”, “50–60 min” ja “Yli 60 min”. Matkojen hinnoille luotiin myös sarake luokittelua varten. Hinnat luokiteltiin kolmeen ryhmään: “Edullisin kolmannes”, “Keskihintainen” sekä “Kallein kolmannes”.

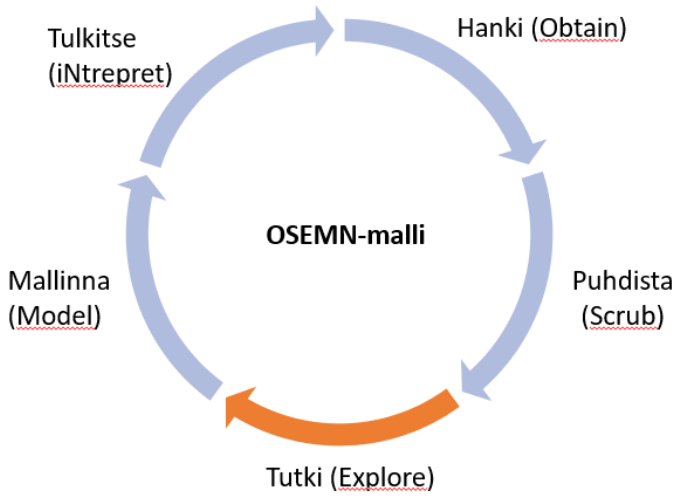
Reklamaatioraportin data

Reklamaatioraportin datan lataaminen Power BI:hin tapahtui noutamalla se lähdekansioon luodusta Excel-työkirjasta. Noudon yhteydessä sarakkeiden tietotyypit muutettiin sopiviksi ja “Reklamaation tila” -sarake poistettiin, koska sen ei katsottu olevan datamallin kannalta merkityksellinen. Muut lähtödatan sarakkeet olivat sellaisia, joita voi mahdollisesti hyödyntää datamallia kehitettäessä, joten niitä ei poistettu. Koska data on asiakaspalvelutiimin manuaalisesti täyttämää ja ylläpitämää niin siitä löytyi joitakin inhimillisiä virheitä kuten päivämäärissä näppäilyvirheitä. Virheiden korjaukset tehtiin Power Queryn sijaan suoraan Excel-työkirjaan, joten Power Queryn puolella tehtyjä muokkauksia tuli varsin vähän ja ne lähinnä olivat termien yhteen mukaistamisia toiseen lähdeaineistoon sopiviksi esimerkiksi kuljetusyhtiöiden nimien osalta.

3.4.5 Datan tutkiminen

Tässä kappaleessa kuvataan datan tutkimiseen käytettyjä toimenpiteitä. Datan tutkiminen on OSEMN-mallin kolmas vaihe (kuva 14). Raportointityökalusta saadun kuljetushallintajärjestelmädatan oikeellisuuden ja virheiden etsimisen tueksi tehtiin Exceliin vastaava datan koonti samoista CSV-tiedostoista. Täten pystyttiin testaamaan ja vertaamaan, että rivimäärät eri suodatuksin ja mittarien antamat luvut ovat paikkaansa pitäviä. Tällä menettelyllä varmistettiin, että datat ovat tarkkoja, asiaankuuluvia sekä ajantasaisia. Datan tutkimisen yhteydessä sille tehtiin puhdistus toimenpiteitä tarvittaessa. Vaikka varsinainen tutkiminen toteutettiin jo työn alkuvaiheessa pienemmällä aineistolla enne varsinaista mallinnusta ja

datamallin rakentamista, niin tutkimisessa käytettyjä toimenpiteitä myös aina silloin kun uutta dataa noudettiin malliin.



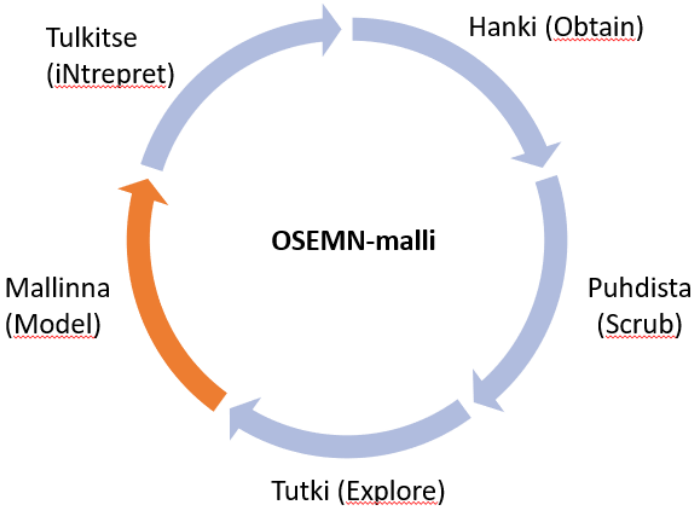
Kuva 14 Datan tutkimisen yhteys OSEMN-malliin

Tutkimis-vaiheessa luotiin DAX-koodikielellä muutamia yksinkertaisia mittareita kuten "Kuljetustilauksia", "Matkustajia" ja "Kohteita". Koska yksi rivi käsittää yhden kuljetustilauksen niin kuljetustilausten määrä saatiin laskettua yksinkertaisesti Countrows -formulalla. Matkustajien ja kohteiden laskemiseen käytettiin formulaa Distinctcount. Nämä mittarit sekä vastaavilla laskuperiaatteilla luodut muut yksinkertaiset mittarit toimivat kehittyneempien laskukaavojen perustana. Datan tutkimiseen käytettiin visualisoinneista hyödyksi muun muassa suodattimia, taulukkoja, kortteja sekä pylväs- ja piirakkadiagrammeja. Niitä ei vielä tässä vaiheessa muokattu asetuksien, kuten värien ja otsikointien, osalta sen enempää vaan pidettiin hyvin pelkistettyinä oletusasetuksineen. Tutkimisvaiheessa saatiin hyvä käsitys käsiteltävän datan laadusta ja sen paikkansapitävyydestä, mikä puolestaan mahdollisti sujuvan siirtymisen seuraavaan vaiheeseen eli mallinnukseen ja datamallin rakentamiseen.

3.4.6 Mallinnus ja Datamallin rakentaminen

Tässä kappaleessa kuvataan datan mallinnuksessa ja datamallin rakentamisessa tehdyt tärkeimmät toimenpiteet. Ensiksi kuvataan datamallin Mallinäkymä, mistä käy ilmi mallinuksessa mukana olevat fakta- ja dimensiotaulut sekä niiden väliset suhteet. Toisena kuvataan valitut päämittarit ja niihin käytetyt laskentamenetelmät. Lopuksi on kuvattu

Raporttinäkymään rakennetut raporttisivut ja niiden visualisoinnit. Mallintaminen on OSEMN-mallin prosessissa neljäs vaihe (Kuva 15).

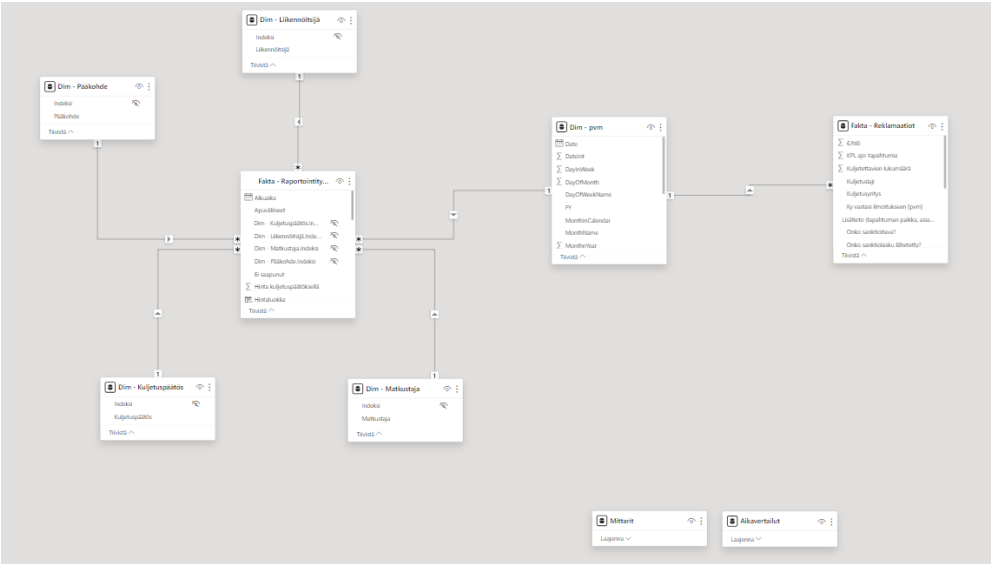


Kuva 15 Mallinnuksen ja datamallin rakentamisen yhteys OSEMN-malliin

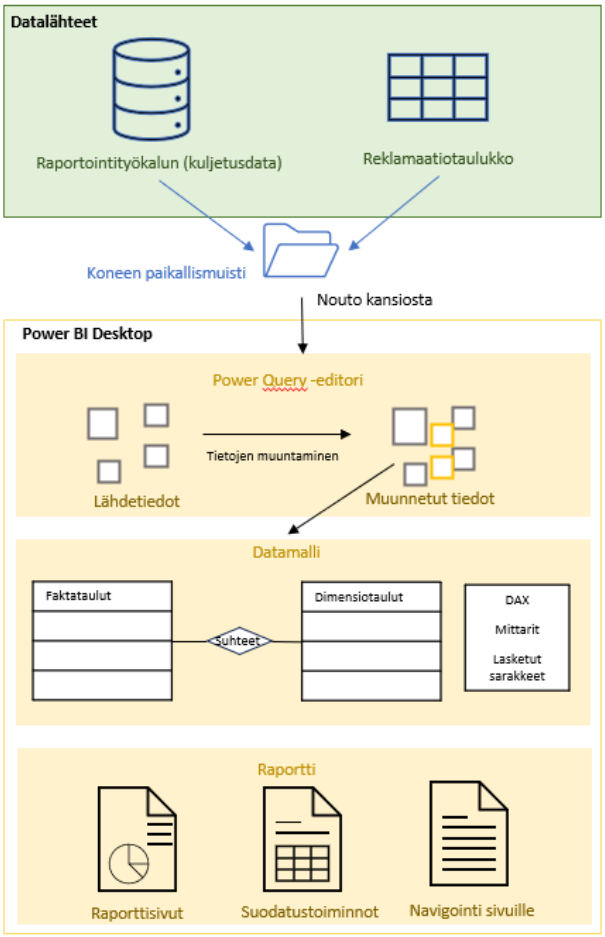
Datamallin rakenne

Datamalli koostuu kahdesta faktataulusta ja viidestä dimensiotaulusta kuten esitetty kuvassa 16. Siinä näkyy myös mallissa mukana olevat mittaritaulut, jotka sisältävät DAX-laskukaavat. Kuten kuvasta huomaa niin mittareilla ei ole suhdetta suoraan datamallin tauluihin. Datamalli on suunnisteltu tähtirakenteen mukaisesti Raportointityökalun datan käsittävän faktataulun ympärille. Dimensiotaulut luotiin sarakkeista Matkustaja, Liikennöitsijä, Pääkohde ja Kuljetuspäätös tekemällä kullekin faktataulusta kaksoiskappale ja poistamalla niistä kaikki muut sarakkeet paitsi tarkoituksen mukainen dimensiosarake. Tämän jälkeen poistettiin kaksoiskappaleet ja lisättiin indeksisarake, jonka jälkeen faktataulu sekä juuri luotu dimensiotaulu yhdistettiin indeksisarakkeen avulla. Faktataulun ja dimensiotaulujen suhteiden kardinaliteetti on Monta yhteen (*:1). Toisin kuin muut dimensiotaulut, niin erillinen päivämäärätaulukko luotiin internet-aineistosta kopioidulla koodilla, missä on laajasti eri sarakkeita ajanjaksoihin liittyen. Sen avulla voitiin tarkastella laajemmin tietoja erilaisilta ajanjaksoilta kuten viikko, kuukausi, kvartaali, vuosi sekä niiden yhdistelmistä eri muodoissa kuten esimerkiksi 2023/02 ja Tammi 2024. Päivämäärädimensiotaulu toimii yhdistävänä

tekijänä kahden faktataulun välillä, eikä muuta yhteyttä luotu niiden kesken tämän työn puitteissa. Datamallin ETL-prosessi on nähtävänä kuvassa 17.



Kuva 16 Datamallin rakenne



Kuva 17 Datamallin ETL-prosessi

Mittarit

Kuten datan tutkimisvaiheessa kuvattiin, kuljetustilaukset saatiin laskemalla yksinkertaisesti rivien määrä, sillä datassa yksi rivi käsittää yhden kuljetustilauksen. Peruttujen kuljetusten lukumäärä saatiin laskettua sarakkeen *peruttu* avulla seuraavan laskukaavan mukaisesti: = *COUNTROWS(FILTER('Fakta - Raportointityökalu', 'Fakta - Raportointityökalu'[peruttu] = TRUE()))*. Kuljetustilauksista kun vähensi tämän peruttujen kuljetusten lukumäärän niin saatiin laskettua saman toteutuneiden kuljetusten lukumäärä. Mikäli kuljetus on peruttu myöhässä niin kuljetus yhtiö voi veloittaa kuljetuksesta. Myöhässä perutut kuljetukset saatiin laskettua sarakkeen *Peruttu myöhässä* avulla seuraavan laskukaavan mukaisesti: = *COUNTROWS(FILTER('Fakta - Raportointityökalu', 'Fakta - Raportointityökalu'[Peruttu myöhässä] = TRUE()))*. Nämä kuljetustilauksiin liittyvät mittarit loivat perustan kuljetusten määrien toteumien tarkastelulle. Näiden avulla myös aikavertailuun liittyviä mittareita

luotiin kuten esimerkiksi kuljetustilausten määrä vastaavana tarkasteluaikana vuosi sitten seuraavan kaavan avulla: *CALCULATE([Kuljetustilauksia], DATEADD('Dim - pvm'[Date], -1, YEAR))*. Tämän avulla voitiin sitten tehdä prosentuaalista vertailua tarkasteluvuoden ja sitä edeltävän vuoden välillä, jonka avulla nähdään ovatko määrät laskeneet vai kasvaneet ja kuinka merkittävästi.

Datamallissa oleellisen mittarina oli myös matkan kesto, mikä saatiin laskettua sarakkeiden *Alkuaika* ja *Päätösaika* avulla seuraavan laskukaavan mukaisesti: = DATEDIFF(MIN ('Fakta - Raportointityökalu'[Alkuaika]), MAX ('Fakta - Raportointityökalu'[Päätösaika]), SECOND). Luettavuuden takia se muokattiin minuuteiksi yksinkertaisesti jakamalla kuudellakymmenellä. Taulukossa 12 on esitetty datamallissa käytetyt tärkeimmät mittarit.

Taulukko 12 Datamallin tärkeimmät mittarit

Mittari	**Laskukaava**	**Selite**
Kuljetustilauksia	= COUNTROWS('Fakta - Raportointityökalu')	Yksirivi käsittää yhden tilatun kuljetuksen, joten kuljetustilausten määrä saadaan laskettua yksinkertaisesti laskemalla rivimäärä.
Toteutuneita Kuljetuksia	= [Kuljetustilauksia] - [Peruttuja kuljetuksia]	Toteutuneet kuljetukset ovat kaikki tilatut kuljetukset joita ei ole peruttu.
Peruttuja kuljetuksia	= COUNTROWS(FILTER('Fakta - Raportointityökalu', 'Fakta - Raportointityökalu'[peruttu] = TRUE()))	Perutut kuljetukset saadaan laskemalla kaikki ne rivit missä peruttu -sarakkeessa on merkintä true.
Myöhässä peruttuja kuljetuksia	= COUNTROWS(FILTER('Fakta - Raportointityökalu', 'Fakta - Raportointityökalu'[Peruttu myöhässä] = TRUE()))	Keskiviikkosäännön mukaisesti peruutukset tehtävä edellisen viikon keskiviikkoon mennessä. Jos peruutus tehty tämän jälkeen eli myöhässä niin kuljetusyhtiö voi veloittaa kuljetustilauksesta. Laskettu jos ”peruttu myöhässä” sarakkeessa on ”kyllä”.

Peruutus %	= [Peruttuja kuljetuksia] / [Kuljetustilauksia]	Peruttujen kuljetusten %-osuus tilatuista kuljetuksista.
Myöhässä peruttujen %	= [Myöhässä peruttuja kuljetuksia] / [Kuljetustilauksia]	Myöhässä peruttujen kuljetusten %-osuus tilatuista kuljetuksista.
Perutuista kuljetuksista myöhässä peruttuja	= [Myöhässä peruttuja kuljetuksia] / [Peruttuja kuljetuksia]	Myöhässä peruttujen kuljetusten %-osuus perutusta kuljetuksista.
Kustannukset	= SUM('Fakta - Raportointityökalu'[Hinta kuljetuspäätöksellä])	Kuljetustilausten kokonaiskustannukset saadaan summaamalla kaikkien rivien hinnat.
Kustannukset toteutuneet kuljetukset	= SUMX(FILTER('Fakta - Raportointityökalu', 'Fakta - Raportointityökalu'[peruttu] = FALSE()),'Fakta - Raportointityökalu'[Hinta kuljetuspäätöksellä])	Toteutuneiden kuljetusten kustannukset saadaan summaamalla kaikkien niiden rivien hinnat missä peruttu sarake on epätosi.
Myöhässä peruttujen kustannukset	= CALCULATE([Kustannukset], (FILTER('Fakta - Raportointityökalu', 'Fakta - Raportointityökalu'[Peruttu myöhässä] = TRUE())))	Keskiviikkosäännön mukaisesti peruutukset tehtävä edellisen viikon keskiviikkoon mennessä. Jos peruutus tehty tämän jälkeen eli myöhässä niin kuljetusyhtiö voi veloittaa kuljetustilauksesta. Laskettu kustannukset silloin jos ”peruttu myöhässä” sarakkeessa on ”kyllä”.
Todelliset kustannukset	= [Kustannukset toteutuneet kuljetukset] + [Myöhässä peruttujen kustannukset]	Toteutuneiden kuljetusten kustannukset + myöhässä peruttujen kuljetusten kustannukset. Sisältää kustannukset mitä kuljetusyhtiöt voivat veloittaa.
Keskihinta	= AVERAGE('Fakta - Raportointityökalu'[Hinta kuljetuspäätöksellä])	Kullakin kuljetuksella on päätös sen hinta. Tämä laskee niiden keskiarvon
Matkustajia	= DISTINCTCOUNT('Fakta - Raportointityökalu'[Matkustaja])	Yksittäisten matkustajien lukumäärä

Kohteita	= DISTINCTCOUNT('Fakta - Raportointityöka-lu'[Pääkohde])	Yksittäisten kohteiden lukumäärä
Kuljetuspää-töksiä	= DISTINCTCOUNT('Dim - Kuljetuspäätös'[Kulje-tuspäätös])	Yksittäisten Kuljetuspäätösten lukumäärä.
Reklamaati-oita	= COUNTROWS('Fakta - Reklamaatiot')	Yksirivi käsittää yhden reklamaa-tion, joten niiden lukumäärä saa-daan laskettua yksinkertaisesti laskemalla rivimäärä.
Reklamaatio %	= [Reklamaatioita] / [Toteutuneita Kuljetuk-sia]	Reklamaatioprosenttia saadaan laskettua
Sanktioita	= COUNTROWS(FIL-TER('Fakta - Reklamaa-tiot', 'Fakta - Rekla-maatiot'[Onko sanktioi-tava?] = "Kyllä"))	Sanktioiden lukumääräsaadaan las-kemalla niiden rivien likumäärä missä “Onko saktioitava?” -sarake on “Kyllä”.
Sanktioiden €	= SUM('Fakta - Rekla-maatiot'[Yhteensä])	Sanktioiden euromäärä saadaan summaamalla sarake Yhteensä
Matkan kesto sek	= DATEDIFF(MIN ('Fakta - Raportointityöka-lu'[Alkuaika]), MAX ('Fakta - Raportointi-työkalu'[Päätösaika]), SECOND)	Matka-ajan kesto sekunteina.
Matkan kesto min	= [Matkan kesto Sek] / 60	Matka-ajan kesto minuutteina.
Matka-ajan ka	= AVERAGE('Fakta - Ra-portointityökalu'[Mi-nuutteja yhteensä])	Matka-ajan kesto keskimäärin.
Yli 60 min kuljetusten määrä	= CALCULATE(COUNT-ROWS('Fakta - Rapor-tointityökalu'), 'Fakta - Raportointityöka-lu'[Matka-aikaluokitus] = "Yli 60 min")	Yli 60 min kestävien kuljetusten lukumäärä saadaan laskemalla nii-den rivien lukumäärä missä Matka-aikaluokitus on “Yli 60 min”.
Ei saapunut	= [Kuljetustilauksia] – COUNTBLANK('Fakta -	Niiden matkojen lukumäärä missä matkustaja ei saapunut. Saadaan

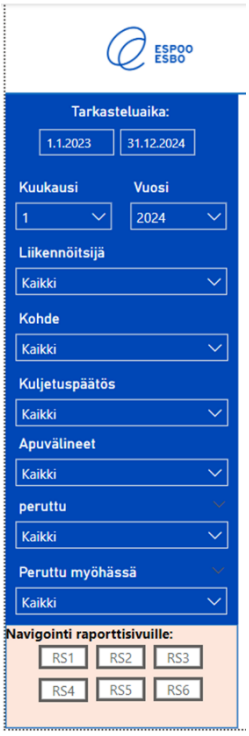
	Raportointityökalu'[Ei saapunut])	laskemalla tyhjät solut sarakkeesta Ei saapunut ja vähentämällä se kuljetustilausten määrästä.
Kuljetustilauksia -1v	CALCULATE([Kuljetustilauksia], DATEADD('Dim - pvm'[Date], -1, YEAR))	Kuljetustilausten lukumäärä vuoden takaa saadaan käyttäen luotua mittaria Kuljestustilauksia, ja yhdistämällä se funktioon DATEADD.
Kuljetustilauksien muutos % -1v	= ([Kuljetustilauksia] - [Kuljetustilauksia -1v]) / [Kuljetustilauksia -1v]	Kuljetustilausten prosentuaalinen muutos vuoden takaiseen voidaan laskea yksinkertaisella jakolaskulla luotujen mittareiden Kuljetustilauksia ja Kuljetustilauksia –1v avulla.
Matkustajia -1v	= CALCULATE([Matkustajia], DATEADD('Dim - pvm'[Date], -1, YEAR))	Matkustajien lukumäärä vuoden takaa saadaan käyttäen luotua mittaria Matkustajia ja yhdistämällä se funktioon DATEADD.

Raporttisivut ja visualisoinnit

Kuljetuspäälliköllä ja Logistiikka-asiantuntijalla on merkittävä rooli yksikön onnistumisessa ja vaikuttavuudessa, joten dataa visualisointiin heidän tarpeidensa mukaisesti. Datamallin visualisoinnit luotiin Power BI Desktopissa kuudelle raporttisivulle, jotka olivat seuraavat: "*1. Johdon näkymä*", "*2. Kuljetusten määrät ja kustannukset*", "*3. Myöhässä peruttujen kuljetusten kustannukset*", "*4. Kuljetusaikojen seuranta*", "*5. Ei saapunut kuljetukseen*" sekä "*6. Kuljetusyhtiöt*". Ensimmäinen sivu on nimensä mukaisesti tarkoitettu johtoryhmälle käytettäväksi ja toinen sivu Logistiikka-asiantuntijalle kuljetusten tarkempaa monitorointia varten. Loput sivuista toimivat kuljetusten laadun seurannan tarkoituksiin. Kukin raporttisivu toimii interaktiivisesti ja raportin käyttäjällä on mahdollisuus muokata sitä esimerkiksi tuomalla malliin uutta dataa tai luomalla uusia mittareita ja visualisointeja. Raporttikokonaisuus on pyritty pitämään kuitenkin kohtuullisena ja selkeänä visualisointien osalta sekä toimivana siten, ettei käyttäjän tarvitse itse muokata mitään. Visualisoinneissa keskityttiin

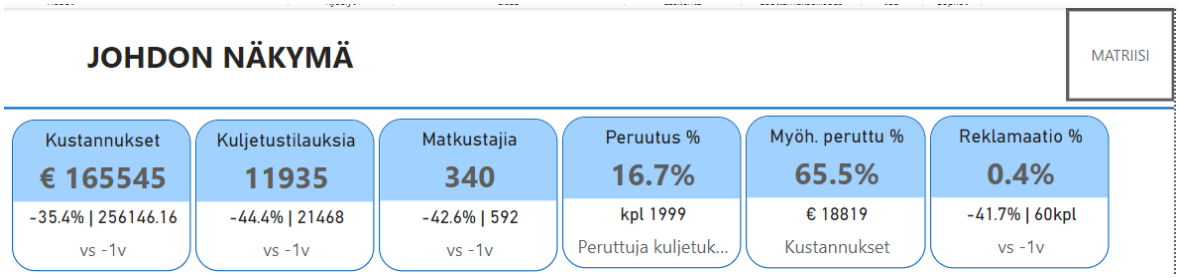
yksinkertaisiin esitystapoihin, joten niissä hyödynnettiin yleisimpiä kaaviotyyppejä viiva- ja pylväsdiagrammeja, ympyräkaavioita, karttoja, kortteja sekä matriiseja.

Kaikissa raporttisivuissa vasemmassa reunassa on samat kuvan 18 mukaiset suodatinmahdollisuudet sekä navigointipainikkeet ohjaamaan halutulle raporttisivulle. Osaan raportin visualisointeihin, missä seurataan mittareita kuukausi- ja vuosikohtaisesti, muotoiltiin vuorovaikutuksia muokkaamalla siten, että kuukausi- ja vuosiosittajilla ei ole vaikutusta. Tämä tehtiin siksi, koska ei haluttu rajata tarkasteluaikaväliä milloinkaan kyseisissä visualisoinneissa. Sivujen ja visualisointien väriteema mukailee pääosin Espoon kaupungin värikoodistoa, mutta myös joissain kohdin väritykseen on valittu sen ulkopuolelta sävyjä. Värikoodisto sekä Espoo logo on otettu aineistopankista, joita voi käyttää Espoon kaupungin esitysaineistoissa. Värikoodiston ulkopuolelta käytettyjä värejä ovat osassa karttanäkymissä käytetyt ympyröiden värit, joissa vihreä symboloi hyvää tilannetta, punainen huonoa tilannetta ja keltainen tilannetta niiden väliltä. Tämä auttaa käyttäjiä nopeasti hahmottamaan tilanteen kriittisyyden.



Kuva 18 Raporttisivun suodattimet ja Navigointipainikkeet

Raporttisivujen ylälaidassa otsikon alla käytettiin KPI-kortteja (kuva 19), jotka luotiin Kortti-visualisointia muotoilemalla. Sitä varten tehtiin tarvittavat mittarit, jotta kortti saatiin kustomoitua haluttuun muotoon.



Kuva 19 KPI -kortit

Kortin yläosaan tuotiin tarkasteltava mittari ja kortin alaosaan mittarista riippuen, vertailu vuoden takaiseen aikaan tai kappale- tai euromäärä. Yläosan mittari piti muuttaa oikeaan formaattiin, jotta sen sai asetettua näkyviin korttiin. Se tapahtui seuraavan esimerkkilasku-kaavan mukaisesti:

```
Kustannukset yht = FORMAT([Kustannukset], "€ #0") & ""
```

Alaosan vertailumittari sen sijaan toteutettiin seuraavan esimerkkilaskukaavan mukaisesti:

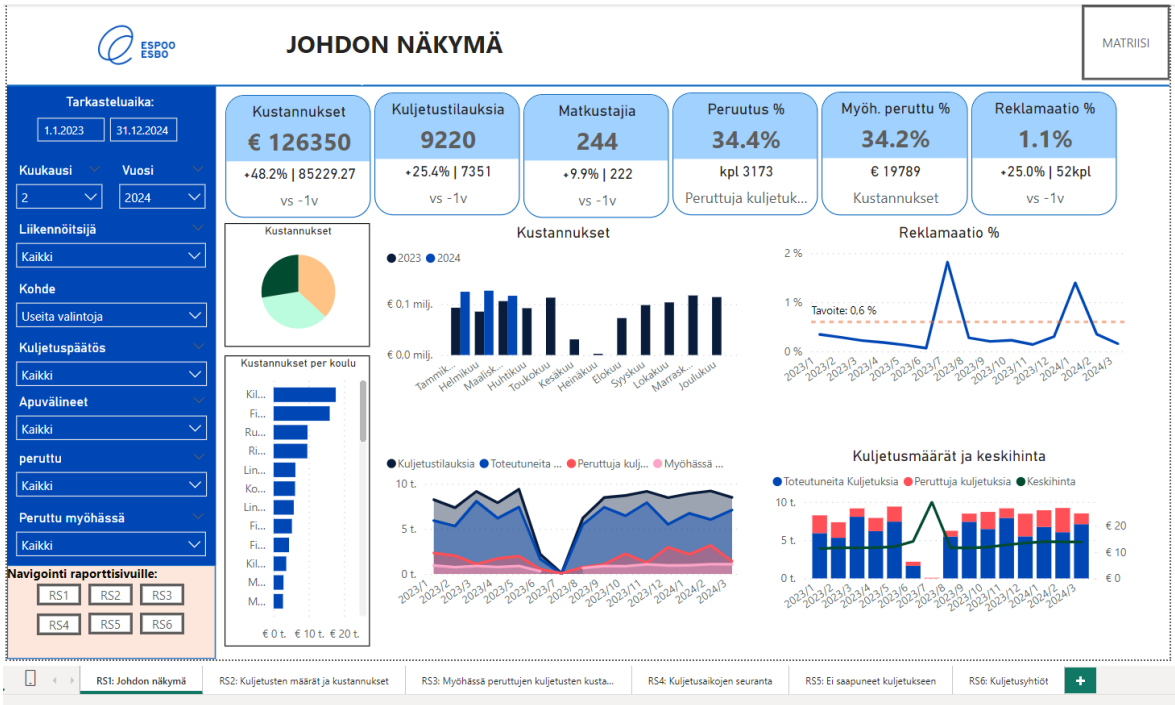
```
Kustannukset vs Kustannukset -1v =
VAR _var = [Kustannukset] - [Kustannukset -1v]
VAR _pct = DIVIDE([Kustannukset], [Kustannukset -1v]) -1
VAR _sign = IF( _var > 0, "+", "" )
RETURN
_sign & FORMAT(_pct, "#0.0%") & " | " & [Kustannukset -1v]
```

Kaikilla raporttisivuilla ei ole käytetty täysin samoja KPI-mittareita vaan valikoiden kunkin sivun kohdalla tapauskohtaisesti. Esimerkiksi reklamaatioprosenttimittari on käytössä vain johdon näkymällä.

Raporttisivu1: Johdon näkymä

Ensimmäisenä raporttisivuna on "*Johdon näkymä*" (kuva 20), jonka tarkoituksena on näyttää olennaiset mittarit ja lukemat helposti johdolle kuukausipalaverissa tai tarvittaessa. Se toimii yleiskatsauksena, jota varsinkin Logistiikkajohtaja ja Kuljetuspäällikkö voivat hyödyntää.

Sen sisältö mukailee osittain jo entuudestaan johdon seuraamia mittareita, kuten kustannukset, matkustajamäärät, kuljetusmäärät ja niiden keskihinta sekä reklamaatioprosentti. Uusina mittareina raporttisivulle tuotiin peruutusprosentti, myöhässä peruttujen prosentti, kustannusten jakautuminen liikennöitsijöiden ja toimipisteiden mukaan.



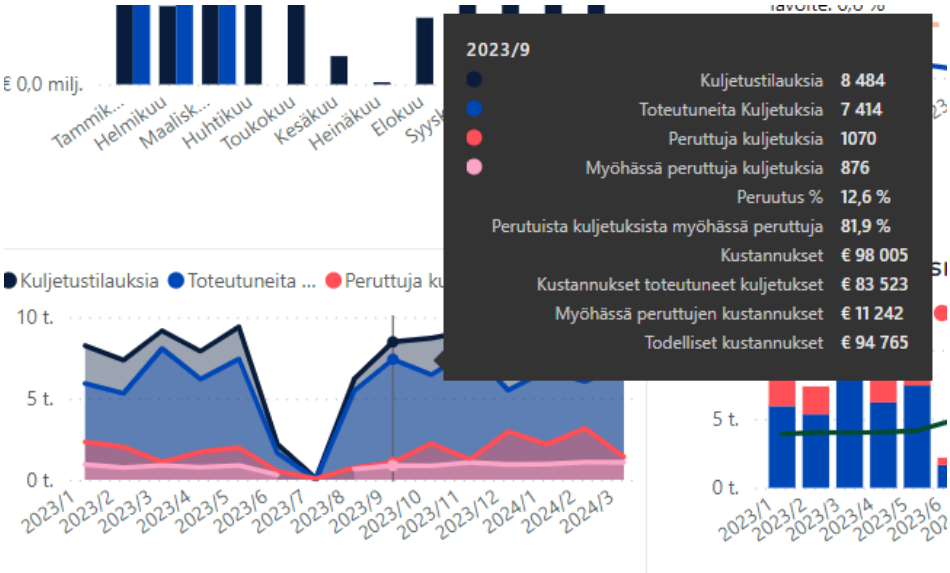
Kuva 20 Johdon Näkymä, datamallin 1. raporttisivu

Kustannuksia liikennöitsijän mukaan visualisoitiin piirakkakaavion muodossa, koska selitteitä oli tarkasteluvälistä riippuen 3–4 kappaletta, mikä on vielä kohtuullinen määrä kyseisessä visualisoinnissa. Toimipisteiden mukaan kustannukset visualisoitiin palkkikaaviolla, jotta saataisiin havainnollistaen pinottua enemmän kustannuksia omaavat kohteet isommiksi palkeiksi ylemmäs ja siitä järjestyksessä laskien alaspäin. Visualisoinnin työkaluvihjeestä pystyy myös tarkastelemaan kohdekohtaisesti kustannukset, matkustajien määrän, kuljetustilausten määrän, toteutuneiden tilausten määrän sekä kuljetusten keskihinnan.

Kustannuksia visualisoitiin klusteroidun palkkikaavion avulla myös kuukausikohtaisesti siten, että selitteenä on vuosi. Koska aineisto kattaa vuoden 2023 ja 2024 dataa niin tästä voidaan verrata, kuinka paljon vuosien välillä on eroa. Lukujen vertaaminen edelliseen kauteen on oleellinen osa toiminnan seurantaa ja sen avulla voidaan ryhtyä selvittämään mahdollisia

poikkeamia heti kun niitä havaitaan. Vertailua tukemaan työkaluvihjeeseen otettiin mittariksi kustannusmuutos vuoden takaiseen, mistä nähdään tarkka prosentuaalinen muutos.

Vertailu kuljetustilauksista (tumman sininen), toteutuneista kuljetuksista (sininen), perutuista kuljetuksista (punainen) sekä myöhässä perutuista kuljetuksista (vaalean punainen) toteutettiin aluekaavion avulla. Vertailusta nähdään kuukausitasolla, paljonko tilatuista kuljetuksista on toteutuneita ja paljonko peruttuja. Kuljetukset tilataan kuljetusyhtiöiltä aina edeltävän viikon keskiviikkoon mennessä ja ne pitää myös perua samaan ajankohtaan mennessä, muuten peruminen merkitään myöhässä perutuksi ja liikennöitsijät voivat veloittaa niistä. Tämän takia tarkasteluun otettiin myös myöhässä perutut kuljetukset erikseen, joiden kustannukset lisäämällä toteutuneiden kuljetusten kustannuksiin saatiin laskettua mittari *Todelliset kustannukset*. Kuvassa 21 näkyy vertailun aluekaaviossa käytetyt työkaluvihjeet mukaan lukien *Todelliset kustannukset*.



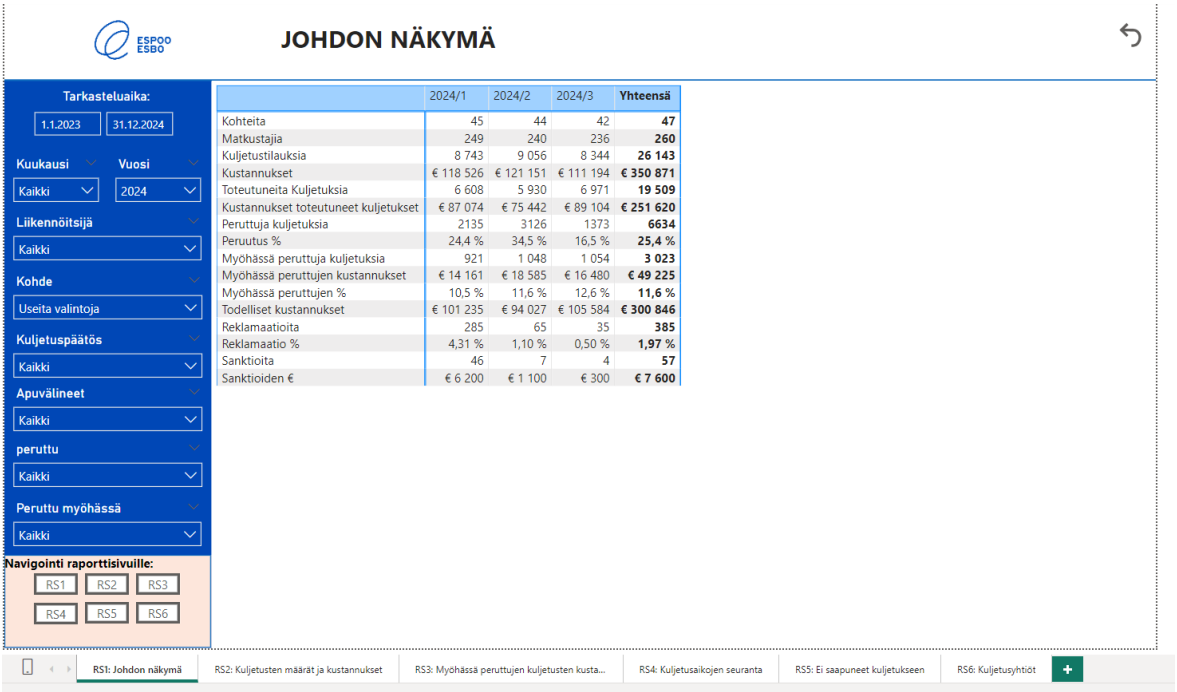
Kuva 21 Työkaluvihjeet aluekaaviossa

Reklamaatioprosenttia, eli reklamaatioiden määrää suhteessa toteutuneisiin kuljetuksiin, seurattiin kuukausitasolla viivakaavion avulla. Tätä on yksikössä seurattu aikaisemmin, mutta koska kyseessä on olennainen kuljetusten laadun seurannan mittari niin se otettiin mukaan myös raporttisivulle johdon näkymään. Visualisointiin asetettiin katkoviivalla 0,6 prosentin tavoitetaso, jonka avulla huomataan tavoitteessa pysyminen ja milloin se on mahdollisesti ylittynyt. Siten voidaan arvioida tavoitetason ylittymisen syitä ja olla tarvittaessa

liikennöitsijöihin yhteydessä lisäselvityksien osalta. Reklamaatioprosentti on valittu myös sivun KPI-mittariksi, mistä näkee tarkastelukuukauden prosentin ja prosentuaalisen muutoksen vuoden takaiseen sekä reklamaatioiden kappalemäärän vuoden takaiselta ajalta.

Kuljetusmäärät ja keskihinta toteutettiin Viivakaavio ja pinottu pylväskaavio -visualisoinnilla. Kuljetusmäärää ja keskihintaa on yksikössä seurattu kuukausitasolla aikaisemmin, mutta uutena ominaisuutena visualisointiin tuotiin peruttujen kuljetusten lukumäärä. Tästä pystytään helposti havaitsemaan peruttujen kuljetuksien määrän suhdetta toteutuneiden kuljetusten määrän. Keskihintaa tarkastellaan koska halutaan havaita, onko kuljetusten hinnoissa ajan myötä merkittäviä muutoksia. Työkaluvihjeeseen asetettiin kustannustietoina toteutuneiden ja myöhässä peruttujen kustannukset sekä näiden summa todelliset kustannukset, mitä voi myös seurata visualisoinnista halutulta kuukaudelta.

Raporttisivun oikeasta ylälaidasta saa avattua myös koko sivun kokoiseksi aukeavan matriisin (kuva 22), missä on vuosi/kuukausi -tasolla eritelty Logistiikalle olennaiset mittarit. Niistä näkee kuljetusten ja matkustajien lukumäärän, kuljetustilausten lukumäärän ja niiden kustannukset, toteutuneiden kuljetusten lukumäärän ja niiden kustannukset, peruttujen kuljetusten lukumäärän, peruutusprosentin, myöhässä peruttujen kuljetusten lukumäärän ja kustannukset, myöhässä peruttujen kuljetusten prosenttiosuuden tilatuista kuljetuksista, todelliset kustannukset eli toteutuneet kuljetukset summattuna myöhässä peruttujen kuljetusten kustannuksilla, reklamaatioiden määrän ja niiden prosenttiosuuden toteutuneista kuljetuksista sekä sanktioiden lukumäärän ja niiden euromäärän.



	2024/1	2024/2	2024/3	Yhteensä
Kohteita	45	44	42	**47**
Matkustajia	249	240	236	**260**
Kuljetustilauksia	8 743	9 056	8 344	**26 143**
Kustannukset	€ 118 526	€ 121 151	€ 111 194	**€ 350 871**
Toteutuneita Kuljetuksia	6 608	5 930	6 971	**19 509**
Kustannukset toteutuneet kuljetukset	€ 87 074	€ 75 442	€ 89 104	**€ 251 620**
Peruttuja kuljetuksia	2135	3126	1373	**6634**
Peruutus %	24,4 %	34,5 %	16,5 %	**25,4 %**
Myöhässä peruttuja kuljetuksia	921	1 048	1 054	**3 023**
Myöhässä peruttujen kustannukset	€ 14 161	€ 18 585	€ 16 480	**€ 49 225**
Myöhässä peruttujen %	10,5 %	11,6 %	12,6 %	**11,6 %**
Todelliset kustannukset	€ 101 235	€ 94 027	€ 105 584	**€ 300 846**
Reklamaatioita	285	65	35	**385**
Reklamaatio %	4,31 %	1,10 %	0,50 %	**1,97 %**
Sanktioita	46	7	4	**57**
Sanktioiden €	€ 6 200	€ 1 100	€ 300	**€ 7 600**

Kuva 22 Matriisi

Raporttisivu	Johdon näkymä
Käyttötapaus	Koulukuljetusten nykytilan arviointi yhdellä silmäyksellä. Seuranta tapahtuu kuukausittaisessa Logistiikan johtoryhmän kokouksessa. Käyttäjinä Logistiikkajohtaja, Kuljetuspäällikkö ja Logistiikka-asiantuntija.
Mittarit ja visualisoinnit	1. Räätälöidyt KPI-kortit (Kustannukset, Kuljetustilauksia, Matkustajia, Peruutus %, Myöh. Peruttu %, Reklamaatio %) 2. Ympyräkaavio kustanuksista liikennöitsijöittäin 3. Klusteroitu palkkikaavio kustannuksista kohteittain 4. Klusteroitu pylväskaavio kustannuksista selitteenä vuosi 5. Aluekaaviolla vertailu kuljetustilauksista, toteutuneista kuljetuksista, perutuista kuljetuksista sekä myöhässä perutuista kuljetuksista 6. Viivakaavio reklamaatioprosentista 7. Viivakaavio ja pinottu pylväskaavio kuljetusmääristä ja keskihinnasta 8. Matriisi valituista mittareista

Raporttisivu2: Kuljetusten määrät ja kustannukset (Asiantuntijan näkymä)

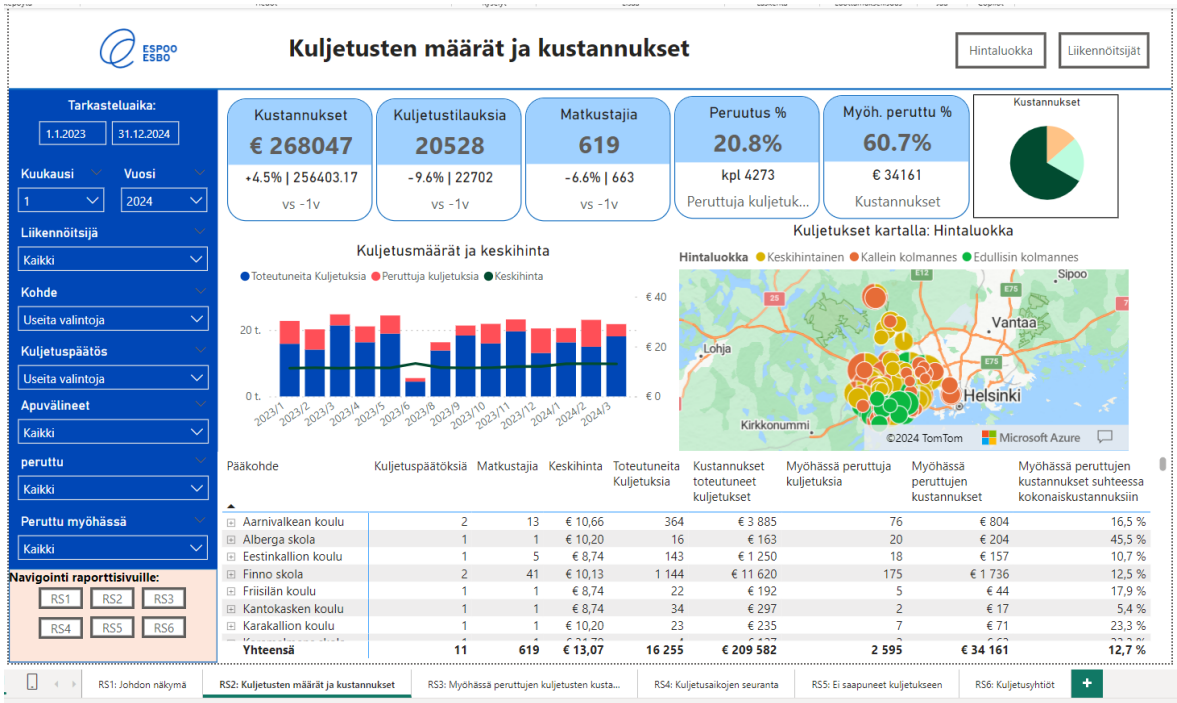
Toisena raporttisivuna on "*Kuljetusten määrät ja kustannukset*" jossa voidaan tarkastella kuljetuksia ja niiden kustannuksia tarkemmalla tasolla kuin johdon näkymässä. Ensisijainen käyttäjä tälle raporttisivulle on Logistiikka-asiantuntija. Käyttäjä voi suodattaa näkymää, jolloin kaikki sivun mittarit muuttuvat vastaamaan valittua suodatusta.

Raporttisivun KPI-kortteina käytettiin kaikkia samoja kuin johdon näkymässäkin pois lukien reklamaatioprosentti, minkä tilalle nostettiin ympyräkaavio kustannuksista liikennöitsijöittäin. Myös visualisointi *Kuljetusmäärät ja keskihinta* on vastaava kuin johdon näkymässäkin oleva. Sen avulla voidaan seurata kuljetusmäärien ja keskihintojen kehitystä, mikä on olennainen tieto myös logistiikka-asiantuntijalle.

Uusina visualisointeina sivulle otettiin Azure-kartta sekä matriisi. Kartan avulla voidaan vaihtaa tarkastelutavaksi koulut *Hintaluokan* (kuva 23) tai *Liikennöitsijän* (kuva 24) mukaan. Painikkeet karttanäkymän vaihtamiseen löytyy sivun oikeasta yläkulmasta. Hintaluokka on jaettu kolmeen osaan siten että, *Edullisin kolmannes* on vihreällä värillä, *Keskihintainen* keltaisella värillä ja *Kallein kolmannes* punaisella värillä. Luokat on luotu lasketulla sarakkeella seuraavaan kaavan mukaisesti:

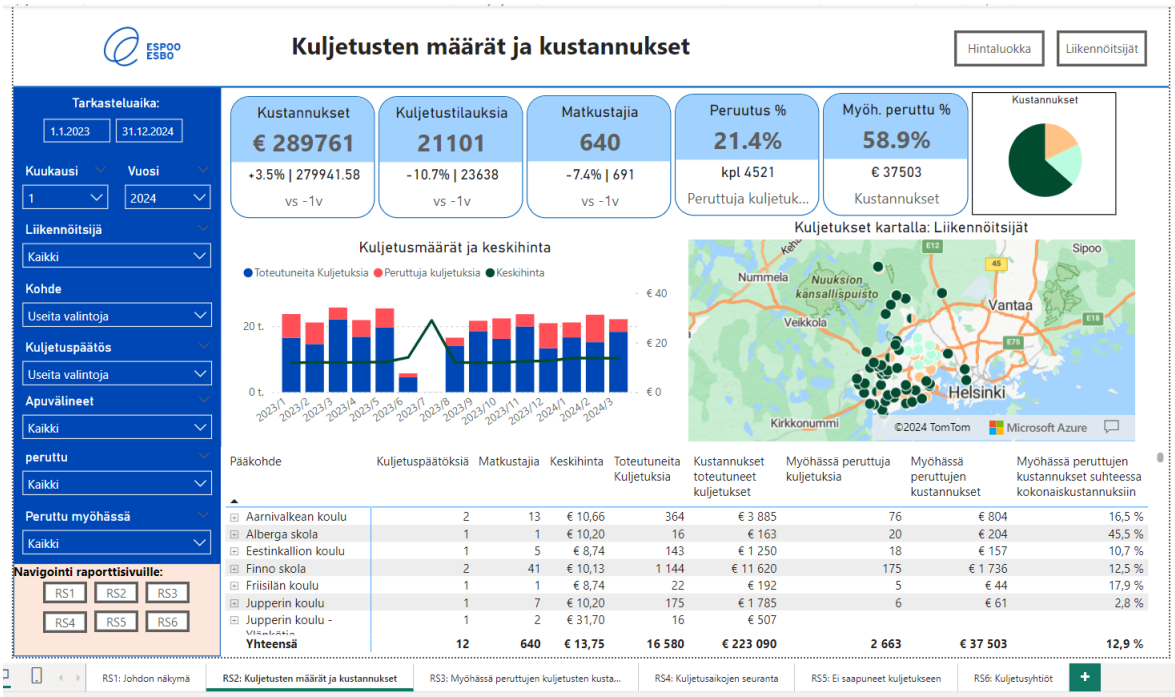
```
Hintaluokka = IF (
        'Fakta - Raportointityökalu'[Hinta kuljetuspäätöksellä] <= PERCENTILE.INC('Fakta
- Raportointityökalu'[Hinta kuljetuspäätöksellä], 0.333),
        "Edullisin kolmannes",
        IF (
            'Fakta - Raportointityökalu'[Hinta kuljetuspäätöksellä] <= PERCEN-
TILE.INC('Fakta - Raportointityökalu'[Hinta kuljetuspäätöksellä], 0.666),
            "Keskihintainen",
            "Kallein kolmannes"
        )
    )
```

Hintaluokkanäkymässä ympyröiden koot mukautuvat toteutuneiden kuljetusten mukaan siten että mitä enemmän kuljetuksia kohteessa on, niin sitä suurempana ympyränä se esiintyy kartalla.



Kuva 23 Kuljetusten määrät ja kustannukset (Hintaluokka -valinnalla), 2. raporttisivu

Kuten kuvassa 24 näkyy niin karttanäkymä liikennöitsijöillä näyttää kohteet sen kuljetusyhtiön värillä mikä ajaa kuljetukset kuhunkin kouluun. Käyttäjä voi hahmottaa eri liikennöitsijöiden toiminta-alueet helposti tämän avulla. Raporttisivun alaosassa olevan matriisin avulla käyttäjä pystyy tarkastelemaan kohdekohtaisesti valittuja mittareita ja tarvittaessa porautumaan siten, että nähdään tiedot kohteen jokaisen matkustajan tarkkuudella. Tarkasteltavana olevina mittareina matriisissa nähdään kuljetuspäätösten ja matkustajien määrä, kuljetusten keskihinta, toteutuneiden kuljetusten määrä ja niiden kustannukset, myöhässä peruttujen kuljetusten määrä ja niiden kustannukset sekä myöhässä peruttujen kustannusten prosentuaalinen osuus kokonaiskustannuksista.



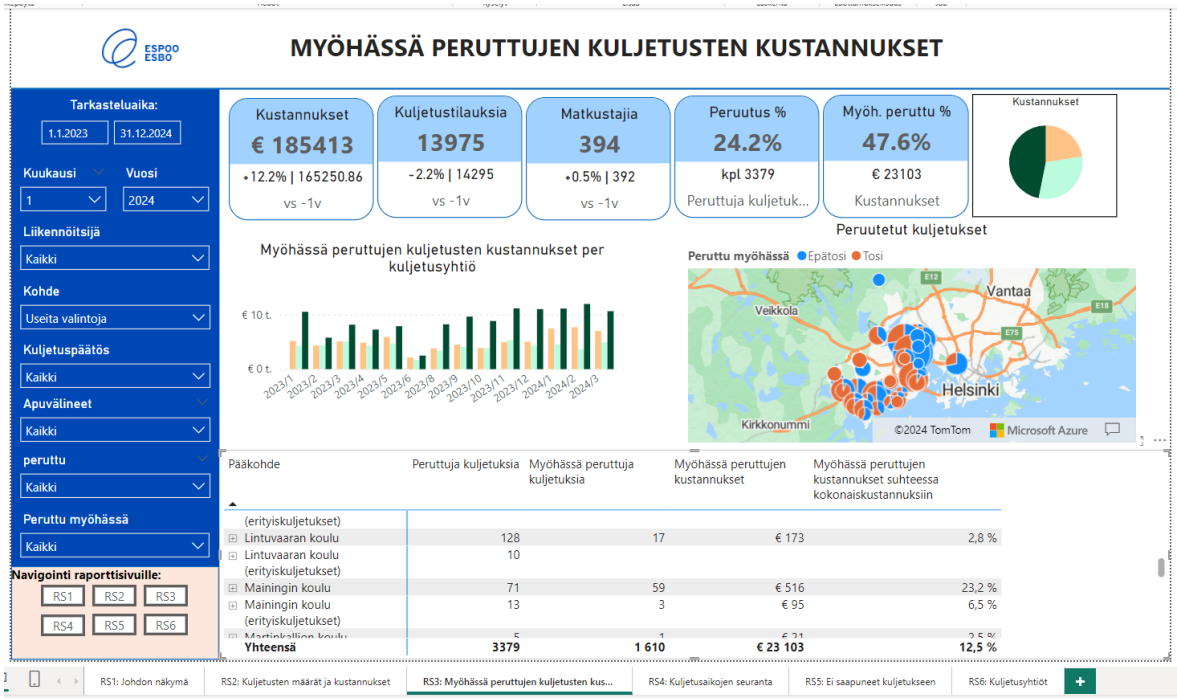
Kuva 24 Kuljetusten määrät ja kustannukset (Liikennöitsijät -valinnalla), 2. raporttisivu

Raporttisivu	Kuljetusten määrät ja kustannukset
Käyttötapaus	Koulukuljetusten tarkempi tarkastelu yksikkötasolla. Mukaan tuotu karttavisuolisoinnit, missä voidaan hahmottaa datan luonne Hintaluokittain ja Liikennöitsijöittäin. Käyttäjänä Logistiikka-asiantuntija.
Mittarit ja visualisoinnit	1. Räätälöidyt KPI-kortit (Kustannukset, Kuljetustilauksia, Matkustajia, Peruutus %, Myöh. Peruttu %) 2. Ympyräkaavio kustannuksista liikennöitsijöittäin 3. Viivakaavio ja pinottu pylväskaavio kuljetusmääristä ja keskihinnasta 4. Azure-kartta kohteista hintaluokan ja liikennöitsijän mukaan 5. Matriisi valituista mittareista kohteittain ja mahdollisuus tarkastella miten ne jakautuvat kohteen kuljetettavien kesken.

Raporttisivu3: Myöhässä peruttujen kuljetusten kustannukset

Kolmas raporttisivu (kuva 25), *Myöhässä peruttujen kuljetusten kustannukset*, on jatkumoa toiselle sivulle ja sen tarkoituksena on tarkastella tarkemmin peruttuja kuljetuksia ja paljonko niistä on peruttu myöhässä, sillä jos peruutus tehdään myöhässä niin kuljetusyhtiö voi veloittaa siitä kuljetuksen hinnan. Sivun ensisijaisena käyttäjänä on Logistiikka-asiantuntija,

joka voi tietojen avulla nähdä, mikäli tietyissä kohteissa tai kuljetettavissa korostuu kuljetusten myöhässä peruminen ja täten tarvittaessa puuttumaan niihin olemalla kouluun tai huoltajaan yhteydessä. KPI-kortteina käytettiin kaikkia samoja kortteja kuin mitkä on esitetty toisen raporttisivun yhteydessä.



Kuva 25 Myöhässä peruttujen kuljetusten kustannukset, 3. raporttisivu

Klusteroidun pylväskaavion avulla tarkastellaan myöhässä peruttujen kuljetusten kustannuksia kuljetusyhtiöittäin kultakin tarkasteluaikavälin kuukaudelta. Tämän avulla voidaan huomata millä kuljetusyhtiöllä korostuu kustannukset myöhässä perumisen takia ja ottaa kuljetukset tarkempaan tarkasteluun raporttisivulla olevan kartan ja matriisin avulla.

Azure-kartalla on visualisoitu kohteet ympyröinä, joiden väritys jakautuu kahteen sektoriin siten, että myöhässä peruttu kuljetus on tosi (oranssilla) tai epätosi (sinisellä). Ympyröiden koko määräytyy myöhässä peruttujen kuljetusten lukumäärän mukaan siten, että mitä enemmän kohteella on peruttuja kuljetuksia niin sitä isompana ympyrä näkyy kartalla. Eli jos esimerkiksi kohteen ympyrä on iso, mutta sininen sektori on iso niin peruttuja kuljetuksia on paljon, mutta ne ovat ajoissa peruttuja. Työkaluvihjeen avulla voidaan myös tarkastella ympyräkohtaisesti matkustajien ja toteutuneiden kuljetusten lukumäärää sekä peruttujen kuljetusten kustannuksia.

Raporttisivun alaosan matriisin avulla tarkastelua voidaan tehdä tarkemmin kohdekohtaisesti. Tarvittaessa voidaan porautua kunkin kohteen kuljetettavien tasolle, mistä sarakkeina toimivien mittareiden tiedot nähdään matkustaja kohtaisesti. Matriisiin valittujen sarakkeiden avulla nähdään peruttujen ja myöhässä peruttujen kuljetusten lukumäärä sekä myöhässä peruttujen kuljetusten kustannukset ja niiden prosentuaalinen osuus kokonaiskustannuksiin nähden.

Raporttisivu	Myöhässä peruttujen kuljetusten kustannukset
Käyttötapaus	Myöhässä peruttujen kuljetusten ja niiden kustannusten tarkkailu ja reagointi, mikäli jossakin kohteessa korostuu voimakkaasti. Käyttäjänä logistiikka-asiantuntija. Asiantuntija voi tarvittaessa ohjeistaa matkustajia tai henkilökuntaa kustannusten vähentämiseksi.
Mittarit ja visualisoinnit	1. Räätälöidyt KPI-kortit (Kustannukset, Kuljetustilauksia, Matkustajia, Peruutus %, Myöh. Peruttu %) 2. Ympyräkaavio kustannuksista liikennöitsijöittäin 3. Klusteroitu pylväskaavio myöhässä peruttujen kuljetusten kustannuksista kuljetusyhtiöittäin 4. Azure-kartta kohteiden perutuista kuljetuksista sen mukaan on se peruttu myöhässä vai ei. Ympyrän koko mukautuu peruttujen kuljetusten lukumäärän mukaan. 5. Matriisi valituista mittareista kohteittain ja mahdollisuus tarkastella miten ne jakautuvat kohteen kuljetettavien kesken.

Raporttisivu4: Kuljetusaikojen seuranta

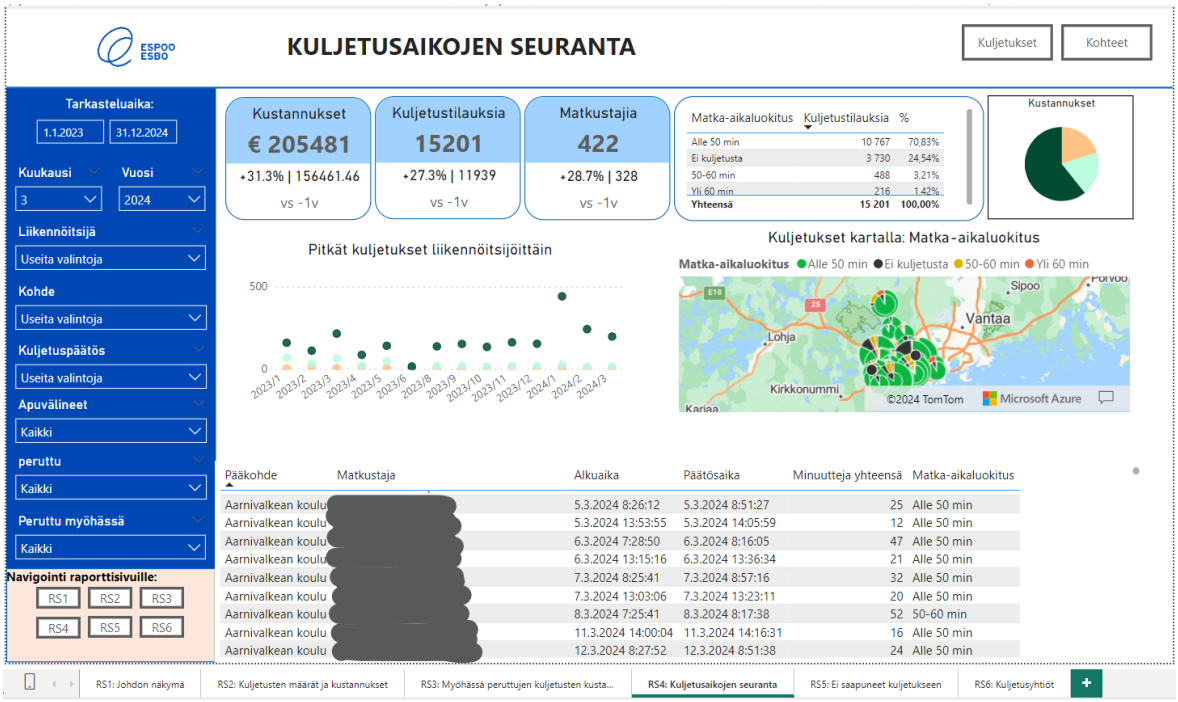
Neljäs raporttisivu, *Kuljetusaikojen seuranta*, toimii kuljetusten laadunvalvonnan tukena. Sen tarkoituksena on, että käyttäjä pystyy seuraamaan koulujen oppilaiden kuljetusaikojen toteutumista sekä kehittymistä läpi tarkasteluajan. Espoossa koulukuljetukset saavat kestää enintään kuusikymmentä minuuttia ja raporttisivun avulla voidaan nähdä halutulta ajanjaksolta kuljetusten aikataulussa pysyminen. Asiantuntija voi olla tarvittaessa kuljetusyhtiöön yhteydessä, mikäli huomataan että kuljetukset ovat kohtuuttoman pitkiä tai jos kuljetuksissa on epäjohdonmukaisuuksia kuten virheellisiä aikaleimoja.

Raporttisivun KPI-kortteina käytettiin mittareita *Kustannukset*, *Kuljetustilauksia* sekä *Matkustajia*. Samalle riville näiden viereen tuotiin matriisi, mistä nähdään Matka-aikaluokitukset ja paljonko kussakin on Kuljetustilauksia sekä prosenttiosuudet suhteessa

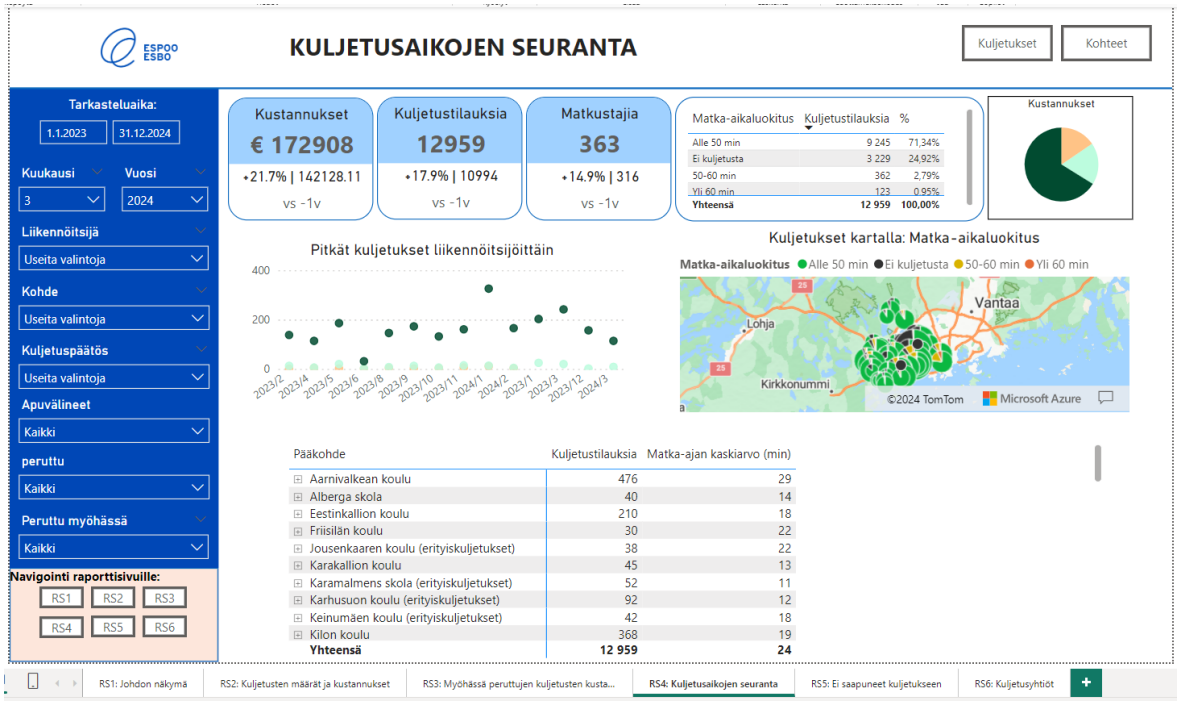
kokonaismäärään. Tämän vieressä on muilta raporttisivuilta tuttu ympyräkaavio kustannuksista liikennöitsijöittäin.

Pistekaavion avulla tarkastellaan pitkiä, yli 60min kestävien kuljetusten määriä liikennöitsijöittäin kuukausitasolla. Tämän avulla voidaan huomata mikäli, jokin kuljetusyrityksistä ajaa tarkasteluaikana paljon ylipitkiä kuljetuksia ja olla heihin yhteydessä sekä pyytää selvitykset tarvittaessa.

Sivulla on kaksi näkymävalintaa, *Kuljetukset* (Kuva 26) ja *Kohteet* (Kuva 27). Ensimmäisessä voidaan tarkastella taulukosta kuljetuskohtaisesti matka-aikoja ja jälkimäisessä voidaan katsoa matriisista kohdekohtaisesti ja tarkemmin myös kunkin kuljetettavan osalta kauanko kuljetukset keskimäärin kestävät.



Kuva 26 Kuljetusaikojen seuranta (Kuljetukset -näkymällä), 4. raporttisivu



Kuva 27 Kuljetusaikojen seuranta, 4 (Kohteet -näkymällä). raporttisivu

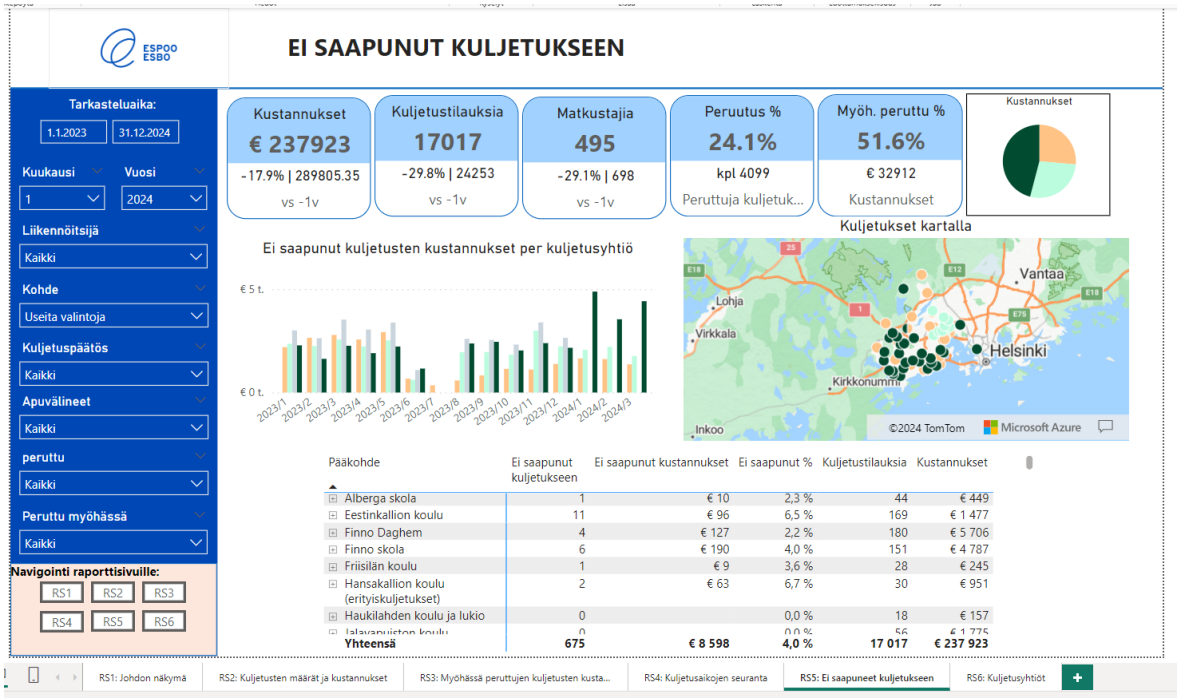
Karttavisualisoinnissa ympyrät ovat kohteita ja niiden suuruus mukautuu toteutuneiden kuljetusten määrän mukaan. Ympyrän väri mukautuu *Matka-aikaluokitusten* mukaan neljään osaan siten että, *Ei kuljetusta* on mustalla värillä, *Alle 50 min* on vihreällä värillä, *50-60 min* on keltaisella värillä ja *Yli 60 min* on punaisella värillä. Luokitukset on luotu lasketulla sarakkeella seuraavaan laskukaavan mukaisesti:

```
Matka-aikaluokitus =
SWITCH(
    TRUE(),
    'Fakta - Raportointityökalu'[Minuutteja yhteensä] <= 0, "Ei kuljetusta",
    AND('Fakta - Raportointityökalu'[Minuutteja yhteensä] > 0, 'Fakta - Raportointityö-
kalu'[Minuutteja yhteensä] < 50), "Alle 50 min",
    AND('Fakta - Raportointityökalu'[Minuutteja yhteensä] >= 50, 'Fakta - Raportointi-
työkalu'[Minuutteja yhteensä] <= 60), "50-60 min",
    'Fakta - Raportointityökalu'[Minuutteja yhteensä] > 60, "Yli 60 min"
)
```

Raporttisivu	Kuljetusaikojen seuranta
Käyttötapaus	Käyttö kuukausittain asiantuntijan työssä. Seurataan kuljetusaikojen toteutumista ja voidaan puuttua virheisiin ja ylipitkiin kuljetuksiin. Käyttäjänä Logistiikka-asiantuntija.
Mittarit ja visualisoinnit	1. Räätälöidyt KPI-kortit (Kustannukset, Kuljetustilauksia, Matkustajia) 2. Matriisi mikä näyttää kuljetustilausten lukumäärän ja prosenttiosuuden Matka-aikaluokituksien mukaisesti 3. Ympyräkaavio kustannuksista liikennöitsijöittäin 4. Pistekaavio, joka näyttää pitkät yli 60min kestävien kuljetusten määrät liikennöitsijöittäin kuukausitasolla 5. Azure-kartta kohteiden kuljetuksista matka-aikaluokituksen mukaan. 6. Taulukko, jossa voi tarkistella kunkin kohteen ja kuljetettavan kuljetustilauksia, niiden aloitus ja päättymisaikoja sekä kestoa ja toteutunutta matka-aikaluokitusta. 7. Matriisi, jossa voi tarkastella kunkin kohteen ja sen kuljetettavien matka-aikojen keskiarvoja.

Raporttisivu5: Ei saapunut kuljetukseen (Laadun valvonta)

Viides raporttisivu (Kuva 28) toimii myös kuljetusten laadunvalvonnan tukena. Sen tarkoituksena on nähdä, että kulkeeko paljon ja korostuuko tiettyjä oppilaita, jotka eivät saavu kuljetukseen. Tällöin asiantuntija voi tarkastuttaa päätökset tekevän toimialan kautta, että onko kuljetustarvetta tai olla kouluun yhteydessä, että peruvat kuljetukset ajoissa, mikäli tiedetään ettei ole kulkemassa kuljetuksilla. Sivun KPI-kortteina käytettiin kaikkia samoja kortteja kuin mitkä on esitetty toisen ja kolmannen raporttisivun yhteydessä.



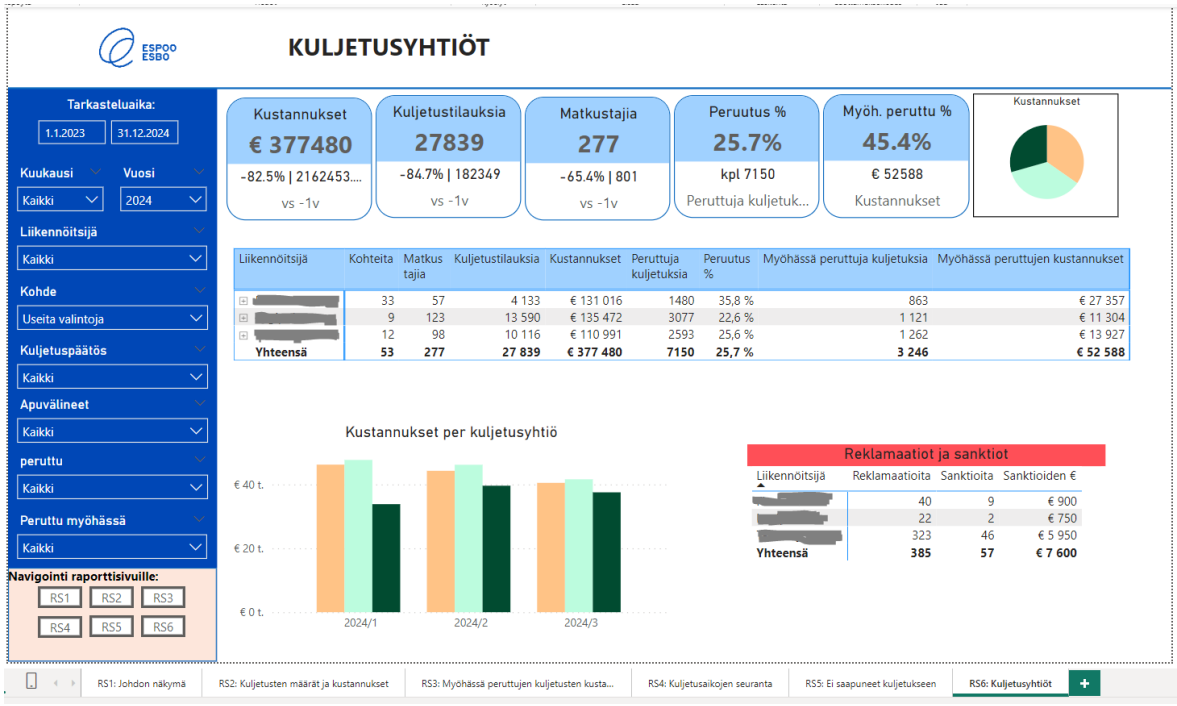
Kuva 28 Ei saapunut kuljetukseen, 5. raporttisivu

Klusteroidun pylväskaavion avulla tarkastellaan niiden kuljetusten kustannuksia kuljetusyhtiöittäin missä kuljetettava ei saapunut kuljetukseen kultakin tarkasteluaikavälin kuukaudelta. Tämän avulla voidaan huomata millä kuljetusyhtiöllä korostuu kustannukset kyseisten syiden takia ja ottaa kuljetukset tarkempaan tarkasteluun raporttisivulla olevan kartan ja matriisin avulla. Matriisin sarakkeiksi valittujen mittareiden avulla nähdään kuljetustilausten lukumäärä ja kustannukset sekä lukumäärää niistä kerroista, jolloin matkustaja ei saapunut kuljetukseen sekä niiden kustannukset ja prosentuaalinen osuus suhteessa kokonaiskustannuksiin.

Raporttisivu	Ei saapunut kuljetukseen
Käyttötapaus	Raporttisivun avulla voidaan seurata niitä kuljetettavia, jotka eivät saavu kuljetukseen ja voidaan löytää kuljetustilauksia (päätöksiä) joiden tarve on lakannut eli niiden kuljetuksia ei käytetä Käyttäjänä Logistiikka-asiantuntija, joka voi ohjeistaa matkustajia ja henkilökuntaa toimimaan sääntöjen mukaan ja peruuttamaan kuljetukset ajoissa, jos tiedetään ettei oppilas ole saapumassa kuljetukseen.
Mittarit ja visualisoinnit	1. Räätälöidyt KPI-kortit (Kustannukset, Kuljetustilauksia, Matkustajia, Peruutus %, Myöh. Peruttu %) 2. Ympyräkaavio kustannuksista liikennöitsijöittäin 3. Klusteroitu pylväskaavio niiden kuljetusten kustannuksista joihin kuljetettava ei saapunut kuljetusyhtiöittäin 4. Azure-kartta kohteiden kuljetuksista liikennöitsijöittäin missä ympyrän koko määräytyy ei saapuneiden lukumäärän mukaan. 5. Matriisi, jossa voi tarkastella kunkin kohteen ja sen kuljetettavien tapauksia missä asiakas ei saapunut kuljetukseen.

Raporttisivu6: Kuljetusyhtiöt

Kuudes raporttisivu (kuva 29) toimii kuljetusyhtiöiden eli sopimustoimittajina toimivien liikennöitsijöiden kuljetusten laadunvalvonnan tukena. Sivun tarkoituksena on olla selkeä kooste kuljetusyhtiöiden tiedoista, joita varsinkin Kuljetuspäällikkö ja Logistiikka-asiantuntija pystyvät hyödyntämään esimerkiksi säännöllisesti järjestettävien sopimusseurantapalaverien valmistelussa. Sivun KPI-kortteina käytettiin kaikkia samoja kortteja kuin mitkä on esitetty toisen raporttisivun yhteydessä.



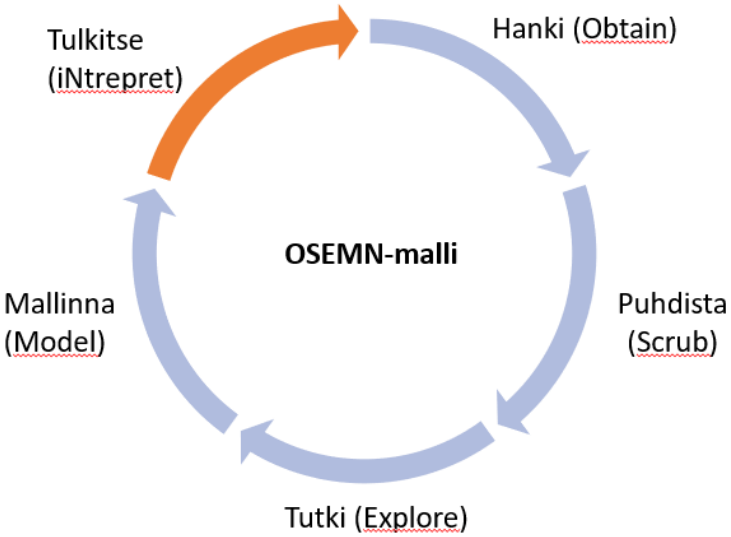
Kuva 29 Kuljetusyhtiöt, 6. raporttisivu

Matriisin avulla pystytään katsomaan kuljetusyhtiöittäin kohteiden tunnusluvut sekä porautua edelleen matkustajien tarkkuudelle toteutumia. Klusteroidun pylväskaavion avulla voidaan verrata kuukausitasolla yhtiöiden kustannuksia. Lisäksi sivuun on valittu erillinen matriisi yhteenvedoksi reklamaatioista sekä sanktioista ja niiden euromääristä yhtiöittäin.

Raporttisivu	Kuljetusyhtiöt
Käyttötapaus	Raporttisivu toimii selkeänä koosteena kuljetusyhtiöiden toiminnasta mitä voidaan hyödyntää sopimusseurantakokouksissa, joissa kuljetusyhtiöiden kanssa käydään toimintaa läpi. Käyttäjänä Kuljetuspäällikkö ja Logistiiika-asiantuntija.
Mittarit ja visualisoinnit	1. Räätälöidyt KPI-kortit (Kustannukset, Kuljetustilauksia, Matkustajia, Peruutus %, Myöh. Peruttu %) 2. Ympyräkaavio kustannuksista liikennöitsijöittäin 3. Matriisi kuljetusyhtiöittäin valituista mittareista. Tarittaessa voi porautua kohteiden ja edelleen niiden matkustajien tarkkuudelle tarkastelemaan sarakkeina toimivia mittareita. 4. Klusteroitu pylväskaavia liikennöitsijöittäin kustannuksista kuukausittain. 5. Matriisi liikennöitsijöittäin reklmaatioiden ja sanktioiden lukumääristä ja paljonko ne on euroissa.

3.4.7 Tulosten tulkinta

Datamalli oli lopulta laajahko kokonaisuus, joka onnistuttiin jakamaan selkeisiin osiin kuudelle raporttisivulle. Datamallin tulosten tulkinta on OSEMN-mallin prosessissa viides vaihe (Kuva 30) ja se toteutettiin yhteistyössä sen ensisijaisten käyttäjien eli kohdeorganisaation Logistiikka-asiantuntijan sekä Kuljetuspäällikön kanssa, joiden työpöydälle tehty datamalli kopioitiin. Tutustuttuaan datamallin sisältämiin ominaisuuksiin he pystyivät antamaan siitä palautetta. Lisäksi he testasivat työkalun avulla meneillään olevan datavarastoprojektin tuottamien tuloksien paikkansa pitävyyttä sekä pystyivät arvioimaan jatkotoimien tarpeellisuutta ja kehityssuuntia. Datamallin laajemman käyttöönoton arviointi ja toteuttaminen tapahtuu tarvittaessa tämän työn ulkopuolella. Datamalli esiteltiin myös Logistiikan johtoryhmän kokouksessa Logistiikkajohtajalle ja muille jäsenille.



Kuva 30 Tulosten tulkinnan yhteys OSEMN-malliin

Ensimmäiselle raporttisivulle, johdon näkymään, oli koottu tärkeimmät tunnusluvut, joiden avulla esimerkiksi Kuljetuspäällikkö voi seurata toimintaa ja laatia arvioita tulevaa kautta varten. Tietoa hyödynnetään muun muassa hankintoja suunnitellessa ja sopimuskauden jatkosta päätettäessä. Graafit toimivat hyvin tiedolla johtamisen tukena ja laajan matriisiin avulla voidaan seurata kätevästi tärkeimpiä koulukuljetuksiin liittyviä lukuja. Lukujen vertaaminen edelliseen kauteen on myös oleellinen osa toiminnan seurantaa ja sen avulla voidaan ryhtyä selvittämään mahdollisia poikkeamia heti kun niitä havaitaan. Lisäksi johdon

näkymään otettu reklamaatioprosentti on laadunvalvonnan tärkein mittari, jota seurataan tarkasti.

Toisen raporttisivun avulla Logistiikkasiantuntija pystyy seuraamaan kustannuksia tarkemmin toimipaikkatasolla ja etsimään poikkeamia sekä tutkia hinnoittelun toteutumista. Visualisoinnit helpottavat käytännön työskentelyä ja ne ovat myös helposti hyödynnettävissä esimerkiksi keskustelun tueksi, jos toimipaikan kanssa suunnitellaan kuljetuksia. Täten myös toimialat pääsevät hyödyntämään tietoja omaa toimintaa, kuten koulupaikkasijoittelua, suunnitellessaan. Sivun molemmat kartat erityisesti helpottavat hinnoittelun ja alueiden seurantaa sekä esimerkiksi mahdollisen uuden aluejaon suunnittelua tulevia kilpailutuksia suunniteltaessa.

Kolmannen raportti sivun avulla Logistiikka-asiantuntija ja johto voivat tehdä seurantaa myöhässä peruttujen kuljetusten kustannuksia sekä helposti löytää haastavia toimipaikkoja, joiden henkilökuntaa voidaan ohjeistaa kustannusten vähentämiseksi. Seuranta on tärkeää nykyisessä hinnoittelumallissa sillä kuljetusyhtiöt voivat veloittaa myöhässä perutuista kuljetuksista.

Neljänteen raporttisivuun oli koottu visualisointeja, joiden avulla Logistiikka-asiantuntija seuraa kuljetusaikoja ja puuttuu rikkeisiin ottamalla yhteyttä kuljetusyhtiöön jopa mahdollisen sanktion muodossa, jos kuljetusajat ylittää 60min. Visualisointeja hyödyntämällä hän pystyy myös näkemään virhemerkintöjä, kuten tavattoman lyhyitä kuljetuksia tai jos esimerkiksi kuljetuksella ei ole päättymisaikaa. Sivun avulla on merkittävästi aiempaa sujuvampaa puuttua ennakoivasti matka-aikoihin, jos ne ovat hyvin lähellä maksimirajaa toistuvasti tietyillä kuljetettavilla.

Viidenteen raporttisivuun oli koottu sellaiset visualisoinnit, joilla voidaan tarkastella kuljetuksia, missä matkustaja ei ole saapunut kuljetukseen. Täten Logistiikka-asiantuntija voi löytää toimipaikkoja, joiden matkustajia ja henkilökuntaa tulee ohjeistaa toimimaan sääntöjen mukaan ja peruuttamaan kuljetukset ajoissa, jos tiedetään ettei oppilas ole saapumassa kuljetukseen. Lisäksi sivun avulla voidaan löytää kuljetustilauksia (kuljetuspäätöksiä) joiden tarve on lakannut eli niiden kuljetuksia ei käytetä. Tätä tietoa voidaan hyödyntää myös uusia päätöksiä tehtäessä.

Kuudennella raporttisivulla on niin sanottu Kuljetusyhtiö -näkymä, mikä on erittäin selkeä kooste kuljetusyhtiön toiminnasta ja toimii loistavasti pohjana esimerkiksi sopimusseuranta

kokouksiin, joissa Kuljetuspäällikkö käy säännöllisesti kuljetusyhtiöiden kanssa toimintaa läpi. Sivulta näkee kätevästi kuljetusyhtiöittäin kohteet ja niiden matkustajat sekä kustannukset, kuljetusmäärät ja reklamaatiot.

Työssä kokonaisuutena rakennettu konstruktio nähtiin kohdeorganisaatiossa onnistuneena. Luotua datamallia pystytään hyödyntämään kuljetusdataan perustuvana analyyttisena työkaluna ja sen visualisointien avulla nähdään datan luonne ja siten pystytään tekemään perustellumpia päätöksiä. Tämä tukee kaupungin strategiaa ja sen tiedolla johtamisen periaatteita. Lopputuotteen myötä raportointi parantuu ja kuljetuslukujen seuranta helpottuu. Luotu tuotos täyttää sille asetetut tavoitteet ja kriteerit, joten se voidaan nähdä onnistuneeksi kokonaisuudeksi. Ratkaisu on kuitenkin vasta ehdotus kokonaisratkaisusta, jota kohdeorganisaatio voi jatkossa kehittää edelleen haluamaansa suuntaan. Havaittuja mahdollisia jatkokehityskohteita ja datamallin laajentamismahdollisuuksia ovat esimerkiksi seuraavat:

- Koodiston luominen sovittamaan kuljetushallintajärjestelmän sekä reklamaatioraportin aineistot yhteen kattavammin.
- Tarkempi seuranta kuljetusyhtiöiden kalustodatasta, mikä on saatavissa järjestelmistä. Kun on nähtävillä heti, että mikä auto kuljetti minkäkin reitin niin voidaan yhdistää reklamaatioihin nopeasti ilman kyselyitä kuljetusyhtiöstä. Kuljetusten täyttöasteen seuranta (Vehicle utilization). Tämän avulla voitaisiin seurata, että taksit eivät kuljeta yli sallittua määrää henkilöitä, mutta toisaalta myös, ettei taksit aja vähällä henkilömäärällä silloin kun kyse ei ole yksilökuljetuksesta.
- Ympäristövaikutukset seuranta ja kaluston sähköistyminen. Voitaisiin seurata täsmäävätkö EU-direktiivien, Hiilineutraali2030 Espoo linjauksen sekä muiden päästötavoitteiden kanssa.
- Kuljetuspäätösdatan analysointi tarkemmin kuljetusten suunnitteluvaiheessa sekä lukuvuoden mittaan. Voidaan nähdä, paljonko kuljetuksia myönnetään kesken lukuvuoden esimerkiksi talviaikaan.

4 Johtopäätökset

Tämän tutkimuksen tavoitteena oli parantaa Espoon kaupungin Logistiikkayksikön hallinnoimien koulukuljetusten toimintaa sekä siihen liittyvää päätöksentekoa tukien kohdeorganisaation tiedolla johtamisen kehittämistä. Työlle oli selkeä tarve, sillä tähän mennessä kohdeyksikössä ongelmana on ollut päätöksenteon tukena olevien mittaristojen ja visualisointien puute. Yksikössä tarvittiin työkalua helpottamaan kuljetusten toteutuksen seurantaa, suunnittelua ja päätöksentekoa erinäisissä työtehtävissä. Tavoitteeseen oli tarkoitus päästä haastattelemalla ensiksi muiden kaupunkien asiantuntijoita ja sen jälkeen rakentamalla analyyttiseksi työkaluksi datamalli haastattelujen tuloksiin sekä verkkoaineistosta löytyvien hyödyllisiksi havaittujen aineistojen ja vastuuyksikön tarpeiden perusteella. Tutkimuksen onnistunut toteutus edellytti tuntemusta kohdeorganisaation vastuuyksikön toiminnasta sekä sen järjestämistä koulukuljetuksista ja niihin liittyvistä asiakokonaisuuksista.

Teemahaastatteluista saaduilla tuloksilla oli tarkoitus vastata toiseen tutkimuskysymykseen: *”Miten Suomen muut kunnat/kaupungit hyödyntävät dataa koulukuljetusten hallinnoinnissa ja mitkä ovat hyviä käytäntöjä, joita Espoon kaupunki voisi ottaa käyttöön toiminnassaan?”*. Haastatteluista huomatiin, että vaikka samat lakiperusteet koskevat kaikkien kaupunkien koulukuljetuksia niin kaupunkien tavoissa järjestää kuljetuksia on eroja. Espoossa, Kuopiossa, Tampereella ja Jyväskylässä datan hyödyntäminen ja tietojen seuranta on eri tasoilla ja intensiteetillä sekä erilaisia työkaluja käytetään datan käsittelyssä. Myös datan käyttökokemus on eri tasoilla, vaikka työtehtävät ovat sisällöiltään enemmän tai vähemmän samankaltaisia keskenään. Kaikista haastatteluista huomattiin, että teknologian ja datan käyttö on keskeinen osa nykyaikaista koulukuljetusalaa, ja työntekijöiden kyky sopeutua niihin liittyviin muutoksiin on tärkeää. Haastateltavat ovat kaikki sitä mieltä, että tulevaisuudessa datan käyttö tulee korostumaan entisestään ja teknologian merkitys tulee kasvamaan, jotta siitä saadaan oleellista tietoa hyödynnetyksi. Paljon vaikuttaa työntekijöiden asenteet ja data itsessään ei siis kannata nähdä pelottavana haasteena vaan pikemminkin mahdollistajana, jotta sitä aidosti pystytään hyödyntämään päätöksenteon tukena. Haastatteluista saatiin Espoolle harkittavaksi se, että otetaan huoltajille sovellus käyttöön kuljetustietojen monitorointia varten. Tämä vähentäisi mahdollisesti yhteydenottaja entisestään, kun vanhemmat näkisivät hyödylliset tiedot suoraan omasta sovelluksestaan. Lisäksi seuraavaa sopimusta

valmisteltaessa mahdollisesti olisi perusteltua muuttaa kuljetusten perumisaikaa lyhyemmäksi nykyisellään mikä toisi kustannussäästöjä. Haastatteluista huomattiin myös korostuvan reittiaikataulujen seurannan tärkeys, mikä vahvisti tarvetta ottaa se mukaan rakennettavaan datamalliin osaksi visualisointeja.

Datamallin ja siitä saatujen tulosten oli tarkoitus vastata ensimmäiseen tutkimuskysymykseen: ”*Miten Espoon kaupunki voi tukea päätöksentekoa Logistiikan koulukuljetusten osalta paremmin dataa hyödyntämällä?*”. Toteutetun datamallin avulla dataa pystyttiin hyödyntämään yksikössä entistä paremmin päätöksenteossa kaupungin strategian tiedolla johtamisen mukaisesti. Mallin kuusi raporttisivua sisältävät monipuolisesti yksikön toiminnan kannalta olennaisia mittareita ja niiden avulla pystyttiin tekemään aikavertailua edeltävään vuoteen sekä katsomaan prosentuaalisia muutoksia. Visualisointien avulla saatavana oleva data pystyttiin laittamaan havainnollistavampaan ja luettavampaan muotoon minkä avulla yksikkö näkee arvokasta tietoa kaupungin koulukuljetuksista. Karttavisualisointien avulla koulut ja oppilaat saatiin tuotua kartalle kuljetusyhtiöittäin ja eri luokituksin. Näiden analyysien avulla kohdeorganisaatio pystyy hyödyntämään kuljetustenhallintajärjestelmää entistä tehokkaammin ja siten tukea kuljetuksiin liittyvää päätöksentekoa. Tietojen perusteella pystytään myös toteuttamaan laadullista seurantaa ja tarvittaessa olemaan yhteydessä asianomaisiin ja ohjaamaan toimintaa tarvittavaan suuntaan. Lisäksi datamallia pystyttiin hyödyntämään laajemman tietovarastohankkeen tukena tulosten testaamisessa, paikkansapitävyyden arvioinnissa sekä kehityssuuntien määrittelyssä.

Tutkimuksen tuloksena syntynyt datamalli on kohdeorganisaation hyödynnettävissä sellaisenaan sisältäen Espoon koulukuljetusdatan vuodelta 2023 ja 2024 kevään osalta. Datamallia on mahdollista kehittää edelleen tulevaisuudessa vastaamaan muuttuvia tarpeita. Lisäksi se on mahdollistaa ottaa laajemmin käyttöön myös muille kuin tässä työssä havaituille ensisijaisille käyttäjille. Vaikka datamallin laajempi käyttöönotto rajautuukin tämän tutkimuksen ulkopuolelle, on kohdeorganisaatiossa koettu tutkimustulokset hyödyllisiksi ja tuotoksesta on tullut hyvää palautetta. Datamalliin tutustuneet käyttäjät kokivat, että se vastasi heidän tarpeisiinsa ja he olivat tyytyväisiä lopputulokseen. Kuitenkin raportin tehokas hyödyntäminen päätöksenteontukena jää kohdeorganisaation vastuulle. Työ toteutettiin konstruktiivisena tapaustutkimuksena ja tutkimustulokset antavat vastauksen organisaation tarpeisiin ja esitettyihin ongelmiin. Näitä tutkimuksen tuloksia voidaan myös hyödyntää muissa konstruktiivisissa tutkimuksissa, sillä tutkimusprosessi noudatti teoreettisen viitekehyksen

mukaisia vaiheita tietotarpeiden määrittämisestä datan mallintamiseen. Tutkimus voisi toimia mallina myös vastaaville tutkimuksille muissa kunnissa tai vastaavan kaltaisilla aloilla. Kuitenkin tämän työn tulosten suora hyödyntäminen sellaisenaan muissa organisaatioissa ei ole mahdollista, mikäli lähtödata poikkeaa tutkimuksessa käytetystä aineistosta.

Jatkotutkimukseksi jäävät aiheet voivat liittyä esimerkiksi kohdeorganisaation tiedolla johtamisen kehittymiseen sekä datamallin vaikutuksiin ja sen laajentamiseen. Havaittuja laajentamisaiheita olivat muun muassa kuljetuspäätösdatan analysointi osana kuljetusten suunnittelua sekä kalustodatan ja kaluston sähköistymisen sekä ympäristövaikutusten seuranta. Lisäksi vastaava haastattelu, kuin mitä tässä tutkimuksessa tehtiin, voitaisiin toteuttaa eri kaupungeille ja kartoittaa heidän tiedolla johtamisen periaatteitaan ja siten saada lisää ideoita toiminnan kehittämiseksi.

5 Yhteenveto

Tietojohtaminen on ilmiönä ajankohtainen niin yksityisellä kuin julkisella sektorillakin. Olennaista on, että saatavilla olevaa dataa pystytään hyödyntämään ja muokkaamaan sellaiseen muotoon, että siitä saadaan hyödynnettävää tietoa päätöksenteon tueksi. Organisaatioiden johtamista toteutetaan siten, että edistetään kykyä luoda arvoa tiedolla ja osaamisella. Pyrkimyksenä on olemassa olevan ja sinne saatavissa olevan tiedon hyödyntäminen organisaatioissa sekä myös saatavuuden varmistaminen jatkossa.

Tässä työssä tutkittiin Espoon kaupungin logistiikkavastuuyksikön tiedolla johtamisen kehittämistä koulukuljetuksiin liittyvän päätöksenteon näkökulmasta. Tavoitteena oli rakentaa datamalli analyyttiseksi työkaluksi minkä avulla kohdeorganisaatio voi hyödyntää saatavana olevaa dataa koulukuljetuksiin liittyvän päätöksenteon tukena. Yksikössä oli aikaisemmin ollut puutetta visualisoinneista, joten datamallin ja sen visualisointien tarkoituksena oli saada data luettavampaan muotoon ja nähdä datan luonne sekä tunnistaa mahdollisia trendejä ja poikkeavuuksia.

Työn tutkimus toteutettiin konstruktiivisena tapaustutkimuksena, missä empiirinen aineisto kerättiin puolistrukturoiduilla teemahaastatteluilla sekä kohdeorganisaation kuljetustenhallintajärjestelmän datasta. Konstruktiona luotiin datamalli mitä kohdeorganisaatio pystyy käyttämään koulukuljetuksiin liittyvän päätöksenteon tukena erinäisissä tehtävissä.

Työn lopputuotoksena saatiin haastatteluista kolmen muun Suomen kunnan näkemyksiä ja kokemuksia esille datan hyödyntämiseen liittyen osana koulukuljetuksia. Lisäksi rakennettu datamallin ja sen visualisointien avulla koulukuljetukseen liittyvä data saatiin nähtävämpään muotoon mistä päätöksentekijät voivat tarvittaessa monitoroida tietoja työtehtävien tukena. Haastattelut toteutettiin loka-marraskuussa 2023 ja datamalli rakennettiin keväällä 2024.

Yhteenvetona voidaan todeta, että opinnäytetyö edustaa tärkeää askelta kohti tiedolla johtamisen parempaa hyödyntämistä koulukuljetusten päätöksenteossa. Sen tulokset voivat toimia esimerkkinä muille kunnille ja organisaatioille, jotka pyrkivät parantamaan päätöksentekonsa laatua ja tehokkuutta datan avulla.

Lähteet

Aigner, A. Miksch S. & Schumann H., 2023 Visualization of Time-Oriented Data. Second edition 2023. London: Springer London

Alpar, Paul, and Michael Schulz. “Self-Service Business Intelligence.” Business & information systems engineering 58.2 (2016): 151–155. Web.

Atlan. (2023) What is Data Modeling? 5 Steps to Building a Robust Model [Verkkoaineisto]. Saatavissa: https://atlan.com/what-is-data-modeling/

Atlan. (2024) Traditional BI vs Self-Serve BI: Which One Suits You the most? [Verkkoaineisto]. Saatavissa: https://atlan.com/self-service-bi-vs-traditional-bi/

Clark, D. (2020) Beginning Microsoft Power BI : A Practical Guide to Self-Service Data Analytics. 3rd ed. 2020. Berkeley, CA: Apress, 2020. Web.

CXToday. (2023) Gartner Magic Quadrant for Analytics and Business Intelligence Platforms 2023 [Verkkoaineisto]. Saatavissa: https://www.cxtoday.com/data-analytics/gartner-magic-quadrant-for-analytics-and-business-intelligence-platforms-2023/

Davenport, T. & Harris, J. 2017. Competing on Analytics: Updated, with a New Introduction: The New Science of Winning

Espoon kaupunki 2023a. Espoo tarina. Luettavissa: https://www.espoo.fi/fi/espoon-kaupunki/espoo-tarina

Espoon kaupunki 2023b. Organisaatio. Luettavissa: https://www.espoo.fi/fi/kaupunki-ja-paatoksenteko/organisaatio

Espoon kaupunki 2023c. Päätöksenteko. Luettavissa: https://www.espoo.fi/fi/espoon-kaupunki/paatoksenteko

Espoon kaupunki 2023d. Espoon kaupungin Intranetsivu Essi. (2023) ABC-tietojohtamiseen. [Verkkoaineisto].

Espoon kaupunki 2023e. Espoon kaupungin sisäinen dokumentti. (2023) Raportin ja analytiikan kokonaisarkkitehtuuri – loppuraportti. [Verkkokaineisto].

Espoon kaupunki 2023f. Espoon perusdiaesitys suomeksi, päivitetty 5.9.2023. Espoon henkilöstön käyttöön.

Espoon kaupunki 2023g. Espoon kaupungin Intranetsivu Essi. (2023) Tiedon hyödyntäminen ja avaaminen. [Verkkoaineisto].

Fink, L., Yogev, N. & Even, A. (2017) Business intelligence and organizational learning: An empirical investigation of value creation processes. Information & Management, Volume 54, Issue 1, 2017, s. 38-56

Goldfedder, J. Building a Data Integration Team: Skills, Requirements, and Solutions for Designing Integrations. 1st ed. Berkeley, CA: Apress L. P, 2020. Web.

Goldstone, R., Pestilli, F. & Börner, K. 2015. Self-portraits of the brain: cognitive science, data visualization, and communicating brain structure and function. Trends in Cognitive Sciences. Vol 19, nro. 8, s. 462-474.

Golfarelli M. & Rizzi S. 2009. Data warehouse design; modern principles and methodologies.

Hirsjärvi, S. & Hurme, H. (2022) Tutkimushaastattelu: teemahaastattelun teoria ja käytäntö. Helsinki, Gaudeamus Helsinki University Press. [E-kirja]. Saatavissa https://www.ellibslibrary.com/book/9789523458123

Hehman, E. & Sally Y. “Doing Better Data Visualization.” Advances in methods and practices in psychological science 4.4 (2021): 251524592110453-. Web.

Helander, N., Ahonen, O., Houhalla, K. & Jääskeläinen, A., Tiedolla johtaminen julkisella sektorilla: käytännön tapauksia eri hallinnon aloilta, Focus Localis, 2023., s. 22–42. Saatavissa: https://journal.fi/focuslocalis/issue/download/7896/1239#page=24

Hotz, N. Data Science Process Alliance. (2023) OSEMN Data Science Life Cycle [Verkkoaineisto]. Saatavissa: https://www.datascience-pm.com/osemn/

IBM. (2024) What is data modeling? [Verkkoaineisto]. Saatavissa: https://www.ibm.com/topics/data-modeling

Janvrin, D. J., Raschke, R. L. & Dilla, W. N. 2014. Making sense of complex data using interactive data visualization. Journal of Accounting Education. Vol. 32, nro. 3, s. 31-48.

Jolly, J. Microsoft Power BI Data Analyst Certification Companion : Preparation for Exam PL-300. 1st ed. 2023. New York, New York: Apress, 2023. Web.

Jääskeläinen, A. & Roitto, J.. “Visualization Techniques Supporting Performance Measurement System Development.” Measuring business excellence 20.2 (2016): 13–25. Web.

Kalita, Jugal K, Dhruba K Bhattacharyya, & Swarup Roy. Fundamentals of Data Science : Theory and Practice. London: Elsevier Science & Technology, 2024. Print.

Kasanen, E., Lukka, K. & Siitonen, A. (1991) Konstruktiivinen tutkimusote liiketaloustieteessä. Liiketaloudellinen aikakausikirja. Vol. 40(3). s. 301–329.

Kimball, R. & Ross, M. The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, John Wiley & Sons, Incorporated, 2013. ProQuest Ebook Central.

Kirk, A., 2019. Data Visualisation: A Handbook for Data Driven Design. 2. painos. London: Sage

Knight, D., Pearson, M., Schacht, B. & Ostrowsky, E. Microsoft Power BI Quick Start Guide: Bring Your Data to Life Through Data Modeling, Visualization, Digital Storytelling, and More, 2nd Edition. 2nd ed. Birmingham: Packt Publishing, Limited, 2020. Print.

Koulukuljetusopas 2019. Espoon kaupunki, Logistiikka 20.6.2019.

KvaliMOTV. 2022. Strukturoitu ja puolistrukturoitu haastattelu. [Verkkodokumentti] Saatavissa: https://www.fsd.tuni.fi/menetelmaopetus/kvali/L6_3_3.html.

Laihonen, H., Hannula, M., Helander, N., Ilvonen, I., Jussila, J., Kukko, M., Kärkkäinen, H., Lönnqvist, A., Myllärniemi, J., Pekkola, S., Virtanen, P., Vuori, V. & Yliniemi, T. 2013. Tietojohtaminen. Tampereen teknillinen yliopisto. [Verkkodokumentti]. Saatavissa: https://trepo.tuni.fi/bitstream/handle/10024/116695/tietojohtaminen.pdf?sequence=2&isAllowed=y

Liew, A. 2007. "Understanding Data, Information, Knowledge And Their Inter-Relationships". Journal of Knowledge Management Practice, Vol. 7, nro.2, s. 1-3.

Listenmaa, Jani. Laita tieto töihin : tiedolla johtamisen käsikirja. Helsinki: Alma Talent, 2023. Print.

Logistiikan organisaation kuvaus power point - LOG Organisaatio 2023. 20.1.2023. Sisäinen materiaali, Espoon kaupunki.

Microsoft. (2023) Mikä on Power BI Desktop? [Verkkoaineisto]. Saatavissa: https://learn.microsoft.com/fi-fi/power-bi/fundamentals/desktop-what-is-desktop

Minh, Q. T., Thai, D. T., Duc, B. T. & Phat, N. H. 2022. “Designing a Data Warehouse Framework for Business Intelligence.” 2022 International Conference on Data Analytics for Business and Industry (ICDABI). s. 498–502.

Myatt, G. & Johnson, W. 2009. Making Sense of Data II a Practical Guide to Data Visualization, Advanced Data Mining Methods, and Applications. 1st edition. Hoboken, N.J: John Wiley & Sons, Print.

Nantasenamat, Chanin. (2020) The Data Science Process [Verkkoaineisto]. Saatavissa: https://resources.experfy.com/bigdata-cloud/the-data-science-process/

Negash, Solomon (2004) "Business Intelligence," Communications of the Association for Information Systems: Vol. 13, Article 15.

Pasian, Beverly. The Constructive Research Approach: Problem Solving for Complex Projects. Surrey, England ; Burlington, Vermont : Gower,. Web.

Pentti Salmela. 2014 Hiljainen tieto, innovaatio ja IT.

Rainardi, Vincent. Building a Data Warehouse : With Examples in SQL Server. Berkeley, CA: Apress, Inc., 2008. Print.

Ranganathan, Kavitha. Impactful Data Visualization : Hide and Seek with Graphs, SAGE Publications India Pvt, Ltd., 2022.

Rowley, J. 2007. "The wisdom hierarchy: representations of the DIKW hierarchy." Journal of Information Science, Vol. 33(2), s. 163-180.

Russo, Marco, and Alberto Ferrari. The Definitive Guide to DAX : Business Intelligence for Microsoft Power BI, SQL Server Analysis Services, and Excel. Second edition. Sydney: Pearson Education, 2020. Print.

Saarteinen, M., Sirenius A. & Skog, L. (2018) Tiedolla johtamisen tila ja kehitysnäkymät valtionhallinnossa 2018 – Asiakastarveselvitys. Valtiokonttori. Saatavissa:

https://cdn.valtiokonttori.fi/wordpress/uploads/sites/4/2019/01/Tiedolla_johtamisen_tila_ja_kehitysnakymat_valtionhallinnossa_2018.pdf

Sabherwal, R. & Becerra-Fernandez I. 2010. Business Intelligence: Practices, Technologies, and Management.

Seamark, Philip. Beginning DAX with Power BI The SQL Pro's Guide to Better Business Intelligence. Berkeley, CA: Apress, 2018. Web.

Seppänen, M. 2023. Tietojohtaminen. [WWW-artikkeli]. Saatavissa: https://www.tuni.fi/fi/tutustu-meihin/tietojohtaminen

Sharma, Vinod, Chandan Maheshkar, and Jeanne Poulose. Analytics Enabled Decision Making. 1st ed. 2023. Singapore: Springer Nature Singapore, 2023. Web.

Sherman, Rick. *Business Intelligence Guidebook : From Data Integration to Analytics*, Elsevier Science & Technology, 2014.

Silverston, L. (2001) The Data Model Resource Book. Volume 1, A Library of Universal Data Models for All Enterprises. Rev. ed. Hoboken, N.J: John Wiley & Sons, Inc., 2001. Print.

Tietojohtaminen-termi. Finto Suomalainen asiasanasto- ontologiapalvelu. Haettu 26.3.2024 osoitteesta https://finto.fi/tt/fi/page/t9

Tilastokeskus. 2023. Väestötietoja maakunnittain: Suurimpien kuntien väkiluku, [Verkkodokumentti]. Saatavissa: https://www.tilastokeskus.fi/tup/suoluk/suoluk_vaesto.html#vaestotietoja-maakunnittain

Unwin, A. (2008) Good Graphics? Handbook of Data Visualization. Berliini, Spinger-Verlag.

Van Der Lans, R. 2012 Data Virtualization for Business Intelligence Systems: Revolutionizing Data Integration for Data Warehouses.

Liite 1. Haastattelun kysymysrunko

	Teema1	Johdanto ja koulukuljetusten järjestäminen
	Teema2	Datan hyödyntäminen, analytiikka ja raportointi organisaatiossa
	Teema3	Kehittäminen ja tulevaisuuden näkymät
	Lopuksi	Lisäykset
T1	1	**Kuvaile rooliasi organisaatiossa ja osallisuudestasi koulutaksiliikenteen hallinnassa?**
T1	2	**Kauanko olet työskennellyt tällä alalla ja millainen on kokemuksesi datan käytöstä?**
T1	3	**Miten Koulukuljetukset on toteutettu kunnassanne?**
T1	4	**Mikä on toimintanne laajuus (koululaisia, kuljetettavia koululaisia, kouluja etc.)?**
T1	5	**Onko toiminta-alue jaettu osa-alueisiin ja montako taksiyritystä operoi alueella?** Miten jako on tehty?
T1	6	**Kuka on vastuussa kuljetusten reitityksestä?** Kunta vai taksiyhtiö?
T2	7	**Millaista dataa organisaationne kerää koulukuljetuksista?**
T2	8	**Miten ja mistä keräätte kyseiset datat?** Kuljetustenhalliintajärjestelmä, manuaalinen syöttö, palautteet? Sisäiset ja ulkoiset tietolähteet?
T2	9	**Minkälaisia haasteita olette kohdanneet datan ja sen hyödyntämisen kanssa?** Esim. virheet, puutteet, ristiriitaisuudet
T2	10	**Miten käsittelette ja analysoitte dataa, jotta saatte siitä jalostettua tietoa?** Mitä digitaalisia työkaluja tai ohjelmistoja käytätte siihen?
T2	11	**Kuinka kehittynyttä analytiikkanne on?** Onko analytiikka kuvailevaa, ennustavaa, ohjailevaa tai autonomista.
T2	12	**Miten raportointi on toteutettu?** Millaisia raportteja tuotatte/seuraatte? Voisiko raportointia parantaa entisestään ja miten?
T2	13	**Mitä mittareita käytätte tai mitä lukuja seuraatte koulukuljetuksiin liittyen?** Kuinka tiheästi lukuja seurataan esim. kerran viikossa tai kuukaudessa? Ketkä lukuja seuraa esim. johtoryhmä, asiantuntija, koordinaattori? Mitä informaatiota haluaisitte nähdä mitä nykyisistä mittareistanne ei välttämättä nähdä?
T2	14	**Mitä visualisointeja käytätte havainnollistamaan minkäkinlaisia tietoja?** Voisiko visualisointeja parantaa entuudestaan? Miten?
T2	15	**Miten hyödynnätte dataa koulukuljetuksiin liittyvässä päätöksenteossa?** Esim. Koulukujetuspäätöksissa, sopismuskilpailutuksiin liittyvän aineiston valmistelussa tai jotenkin muuten?
T2	16	**Onko teillä tapausta, jossa datan avulla saadut oivallukset johtivat merkittävään parannukseen/muutokseen toiminnassanne?**
T2	17	**Millaista tietoa raportoitte sidosryhmille (esim. opetuksen toimiala, vanhemmat tai kuljetusyhtiö)?**
T2	18	**Käytetäänkö reaaliaikaisen tiedon jakamiseen vanhemmille tai kouluille erityisiä alustoja tai sovelluksia?**
T3	19	**Miten näette datan roolin kehittyvän koulukuljetusten hallinnassa seuraavien vuosien aikana?**
T3	20	**Onko olemassa nousevia teknologioita tai trendejä, joita seuraatte ja jotka saattavat vaikuttaa siihen, miten dataa käytetään alalla tai voisivat edelleen parantaa kuljetusten tehokkuutta?**
T3	21	**Onko olemassa täyttämättömiä tarpeita tai kohteita, joilla näette mahdollisuuksia tulevaisuuden dataperusteisille innovaatioille?**
L	22	**Onko teillä lisättävää datan roolista toiminnassanne?**
L	23	**Onko teillä suosituksia muille organisaatioille, jotka haluavat parantaa dataan liittyviä käytäntöjään koulukuljetuksiin liittyen tai muuten?**