

Satu Kangas

LIIKETOIMINTATIEDON HALLINNAN TYÖKALUJEN SOVELTUVUUDEN VERTAILU

Opinnäytetyö

Liiketalouden ammattikorkeakoulututkinto

Tietojenkäsittely

2024



**Kaakkois-Suomen
ammattikorkeakoulu**

Tutkintonimike	Tradenomi (AMK)
Tekijä/Tekijät	Satu Kangas
Työn nimi	Liiketoimintatiedon hallinnan työkalujen soveltuvuuden vertailu
Toimeksiantaja	Neste Oyj
Vuosi	2024
Sivut	42 sivua
Työn ohjaaja(t)	Heli Manninen

TIIVISTELMÄ

Opinnäytetyön tavoitteena oli tutustua liiketoimintatiedon hallinnan peruskäsitteisiin ja tehdä tuotevertailu liiketoimintatiedon hallinnan työkalujen välillä. Tuotevertailun lopputulemaksi tavoiteltiin yhteenvertausta siitä, miten hyvin vertailtavat tuotteet vastaavat toimeksiantajan tarpeisiin. Toimeksiantajan pyynnöstä tehtiin tuotevertailu Power BI -työkalun ja QlikView'n välillä. Työn tutkimusongelmina käsiteltiin, mitä liiketoimintatiedon hallinta tarkoittaa ja millaisia työkaluja Power BI ja QlikView ovat. Tämän lisäksi työssä oli tarkoitus selvittää, miten tuotevertailun kriteerit valitaan. Kirjallisuuskatsaus vastaa näihin kysymyksiin, ja näitä tietoja hyödynnettiin työn käytännön osuudessa.

Työn käytännön osuudessa on esitetty tuotevertailuun valitut kriteerit painoarvoineen. Kriteerit valittiin yhdessä toimeksiantajan eri tehtävissä toimivien edustajien kanssa. Kirjallisuuskatsauksen aikana esille nousseet toiminnallisuudet luokiteltiin toiminnallisuutta kuvaaviin kategorioihin. Tärkeimmiksi valintakriteereiksi valittiin käyttäjäkokemus, raporttien kehitettävyyden, tiedon jakelu, tietoturva ja kustannukset. Käytössä olevia raportteja takaisinmallinnettiin Power BI -raporttikokonaisuudeksi käyttäjäkokemuksen arviointia varten. Lopulta työkalut arvosteltiin arvosanoilla, jotka annettiin yhdessä toimeksiantajan edustajien kanssa kriteerien tärkeyden ja toteutumisen perusteella. Työssä on esitetty kooste tuotevertailun tuloksista.

Opinnäytetyössä tehtiin tuotevertailu, jonka tuloksena saatiin numeerinen arvosana Power BI:lle ja QlikView'lle. Arvioinnin tuloksena voitiin todeta, että Power BI on toimeksiantajalle soveltuvampi työkalu. Arvioinnin lisäksi toteutettiin Power BI -raportti, jota voidaan jatkokehittää.

Asiasanat: liiketoimintatiedon hallinta, datan visualisointi, tuotevertailu, Power BI, QlikView

Degree title	Bachelor of Business Administration
Author (authors)	Satu Kangas
Thesis title	A suitability assessment of business intelligence tools
Commissioned by	Neste Oyj
Time	2024
Pages	42 pages
Supervisor	Heli Manninen

ABSTRACT

The goal of this thesis was to get familiar with the basics of business intelligence, and finally complete a suitability assessment which included the chosen business intelligence tools. The purpose of the suitability assessment was to summarize how assessed business intelligence tools would meet the needs of the thesis commissioner. The tools chosen for the suitability assessment were Power BI and QlikView. The research problems were what business intelligence meant, what kind of tools Power BI and QlikView were, and how to choose assessment criteria and ways of working for a suitability assessment. The literature survey answered these questions, and the information was also used in the practical implementation of this thesis.

In the practical implementation of this thesis, the chosen criteria for suitability assessment were presented with their given scores. The criteria were chosen together with end users in different positions. The properties of the tools were categorized. User experience, report development, distribution of information, information security and expenses were chosen as the most important assessment criteria. Existing reports were also reversely engineered into Power BI to assess user experience between systems. Finally, the properties of the tools were scored together with the commissioner, depending on how well they matched the assessment criteria. The results of the suitability assessment were presented in the thesis.

A suitability assessment of Power BI and QlikView was done, which resulted in a numeric grade for both tools. It can be concluded that Power BI is more suitable for the commissioner than QlikView. In addition to the suitability assessment, a Power BI report was developed. The report can be developed further.

Keywords: business intelligence, data visualization, suitability assessment, Power BI, QlikView

SISÄLLYS

1	JOHDANTO	5
2	LIIKETOIMINTATIEDON HALLINTA ELI BUSINESS INTELLIGENCE.....	5
3	DATAN MATKA TIEDOKSI	6
4	BI-TYÖKALUT	8
4.1	Tietomallit BI-työkaluissa	9
4.2	Datan visualisointi	11
4.3	Power BI.....	12
4.3.1	Power BI Desktop	13
4.3.2	Power BI -palvelu ja Power BI Report Server	19
4.3.3	Lisensointi	22
4.4	QlikView	23
4.4.1	QlikView-tuoteperhe ja arkkitehtuuri	24
4.4.2	QlikView-dokumenttien kehittäminen	25
4.4.3	Lisensointi	29
5	TUOTEVERTAILUN VAATIMUKSET	30
6	TUOTEVERTAILUN TOTEUTUS	34
7	PÄÄTÄNTÖ.....	39
	LÄHTEET	41

1 JOHDANTO

Opinnäytetyön tavoitteena oli tutustua liiketoimintatiedon hallinnan perusteisiin ja selvittää, miten onnistunut tuotevertailu tehdään valittujen työkalujen välillä. Tuotevertailu tehtiin kahden liiketoimintatiedon hallinnan työkalun välillä toimeksiantajan pyynnöstä. Valitut työkalut olivat Power BI ja QlikView. QlikView on ollut toimeksiantajan käytössä rinnakkain Power BI:n kanssa. Tuotevertailun tuloksia käytetään perusteluina sille, voidaanko QlikView’stä luopua. Raporttien laatiminen Power BI:lle vaatii muun muassa kehitysresursseja ja käyttäjien koulutusta, minkä vuoksi olemassa olevien QlikView-dokumenttien siirtäminen Power BI:lle tulee olla perusteltua.

Opinnäytetyössä käydään läpi liiketoimintatiedon hallinnan peruskäsitteitä eli mitä se on ja mikä on sen tarkoitus. Opinnäytetyössä kerrotaan, miten datasta muodostetaan merkityksellistä tietoa datankäsittelyprosessissa. Liiketoimintatiedon hallinnan käsitteitä käsitellään yleisesti ennen vertailtavien työkalujen esittelyä. Lopulta kirjallisuuskatsauksessa kerrotaan, miten onnistunut tuotevertailu tehdään.

Tuotevertailuun valittiin vaatimuksia toimeksiantajan tarpeiden mukaan. Vaatimuksille annettiin painoarvo sen mukaan, kuinka tärkeää on tarpeen täyttyminen. Tuotevertailua varten tehtiin Power BI -raportti, jota vertailtiin samaa dataa käyttävään QlikView-dokumenttiin.

Työn tulokseksi saatiin Power BI -raportti ja yhteenveto työkalujen ominaisuuksista. Arvostelussa esitetty työkalujen numeerinen arvosana ja yhteenvedossa sanallinen arviointi.

2 LIIKETOIMINTATIEDON HALLINTA ELI BUSINESS INTELLIGENCE

Liiketoimintatiedon hallinta (engl. business intelligence tai business insights, myöhemmin BI) on kattotermi erilaisille työvaiheille ja työkaluille, joiden avulla datasta eli tietoaineistoista tuotetaan yrityksen toiminnan kannalta merkityksellistä informaatiota. BI:n tavoitteena on tukea päätöksentekoa ja pienentää siihen liittyviä riskejä. Organisaatio voi saada merkittävän edun kilpailijoihinsa nähden, jos liiketoimintatieto on oikea-aikaisesti käytettävissä. Oikea-aikai-

suus tarkoittaa sitä, että esimerkiksi informaatio on saatavilla ihanteelliseen aikaan päätöksentekoa varten. Organisaatio voi hyödyntää liiketoimintatietoa monilla eri tavoilla. Sen avulla voidaan ennustaa, minkälaisiin lopputuloksiin strategiset päätökset voivat johtaa tai parantaa operatiivista toimintaa analysoimalla menneiden päätösten vaikutusta nykyhetkeen. Päätöksentekoprosessista voidaan tehdä kokonaan tai osittain automaattista käyttämällä BI-työkaluja. Päätöksenteon lisäksi informaatiota voidaan käyttää tiimien toiminnan ja työympäristön parantamiseksi. Yksittäiset työntekijät voivat seurata esimerkiksi oman toimialueensa suoriutumista erilaisten raporttien ja sovellusten avulla. (Loshin 2012, 92–93, 196; Maheshwari 2014, 3, 21–24.)

Datan käsittelyyn ja analysointiin kuluu paljon aikaa ja työtä, minkä vuoksi ongelmien ratkaisemiseen BI:n tekniikoiden ja työkalujen avulla tulisi tuottaa liiketoiminnallista arvoa, jotta ratkaisuihin on perusteltua käyttää resursseja. BI:n arvo muodostuu muun muassa seuraavista tekijöistä: taloudellinen arvo, kun kustannukset pienenevät, tuottavuuden ja voiton kasvu, luottamuksen kasvu ja riskien pieneneminen. (Loshin 2012, 17; Maheshwari 2014, 26; Sherman 2014, 27.)

BI:n tärkeimpiä ominaisuuksia on kyky mitata asioita datan avulla. Organisaation toimintaa voidaan seurata mittareilla (engl. measure). Tätä mittareiden lukuarvoa verrataan johonkin tavoitearvoon. Organisaation toiminnalle kriittistä suorituskykymittaria kutsutaan KPI:ksi (engl. key performance indicator). Usean KPI:n seuraaminen muodostaa laajan kuvan organisaation suoriutumisesta suhteessa tavoitteisiin. Mittari voidaan esittää visuaalisesti esimerkiksi kuvaajana, jossa on edistymispalkki sekä tavoitearvo viivana, jota kohti palkki etenee. Se voi olla myös tunnusluku kuten esimerkiksi lukumäärä, keskiarvo tai muu aggregaatti. (Loshin 2012, 18–20, 94; Knight ym. 2018, 125.)

3 DATAN MATKA TIEDOKSI

Kaikkea tallennettavissa olevaa tietoa kutsutaan dataksi. Data voi olla erilaisia tietotyyppisiä, kuten tekstiä, numeroita tai kuvia. Data voi olla jäsenneltyä tai jäsentämätöntä dataa, jonka rakenne on vaihtelevaa. Datalle on tunnusomaista se, että sitä ei ole vielä käsitelty. Voidaan sanoa, että datalla ei ole ar-

voa sellaisenaan. Organisaatiot tuottavat toiminnoissaan paljon erilaista dataa. Datasta muodostuu informaatiota tiedonkäsittelyprosessin aikana, jossa sitä ensin kerätään, sitten muokataan ja luokitellaan. Prosessissa datalle voidaan tehdä laskelmia tai sitä voidaan ryhmitellä valittujen kriteerien mukaan. Informaation avulla voidaan ymmärtää ilmiöitä, joista data on kerätty. Useita eri tietolähteitä voidaan yhdistää uuden merkityksellisten tiedon (engl. insight) muodostamiseksi. (Loshin 2012, 8–9; Maheshwari 2014, 7–8; Sherman 2014, 8–10.)

Dataa kuvailevaa tietoa kutsutaan metatiedoksi. Metatieto kertoo mitä datalle on tehty, miten sitä on käytetty ja miten sitä tullaan käyttämään. Metatieto voidaan luokitella kahteen luokkaan: tekniseen metatietoon eli esimerkiksi mitä tietotyyppiä data on, tai millainen on tietokentän merkkimäärän yläraja. Metatieto voi olla myös liiketoiminnallista, jolloin se kuvaa datan merkityksen liiketoiminnan kannalta. Datasta pitäisi olla tiedossa ainakin määritelmät, tietolähteet ja mitkä sovellukset sitä käyttävät. (Loshin 2012, 8–9; Maheshwari 2014, 7–8; Sherman 2014, 8–10, 79, 99.)

Analysointiin ja raportointiin tarvittava data voidaan tallentaa tietovarastoon. Tietovarastoon tallennettavan datan avulla täytyy pystyä vastaamaan organisaatiolle tärkeisiin kysymyksiin, ja sen avulla tulisi pystyä muodostamaan kokonaisvaltainen kuva tarkasteltavana olevasta aiheesta. Tietovaraston tulisi olla käytettävissä ilman käyttökatkoksia, ja datan tulisi olla helposti noudettavissa erilaisilla työkaluilla analysointia varten. Tietovarastoon voidaan tallentaa dataa eri tietolähteistä ja samalla operatiivisten tietojärjestelmien kuormitus pienenee, sillä tietokyselyt voidaan tehdä yhteen kohteeseen usean eri järjestelmän asemesta. (Maheshwari 2014, 10–12, 37–39, 42; Sherman 2014, 14.)

Datan yksityiskohtaisuutta kutsutaan rakeisuudeksi. Operatiivisiin tietokantoihin tallennettava data on usein paljon yksityiskohtaisempaa kuin mitä raportointiin tarvitaan. Data tallennetaan tietovarastoon tarvittavalla rakeisuudella. Toisin sanoen tapahtumat voidaan tallentaa minuuttitasolla, mutta raportointiin tarvitaan päivätasolla oleva tieto. Tietovarastoissa olevan datan määrä kasvaa ajan kuluessa, sillä uusi data täydentää olemassa olevaa dataa. Tämä mahdollistaa pitkän ajan trendien muodostamisen. Tietovaraston data ei saa olla

haihtuvaa, eli se tallennetaan tietovarastoon pysyvästi. (Maheshwari 2014, 10–12, 37–39, 42; Sherman 2014, 11–12.)

BI-työkalujen taakse piiloutuva tiedonkäsittelyprosessi on työläs käsin tehtäväksi, mutta prosessia voidaan automatisoida siihen tarkoitettujen ETL-työkalujen (engl. extract-transform-load) avulla. Data vietään tietovarastoon ETL-prosessissa, joka koostuu useista vaiheista. Data haetaan tietojärjestelmistä esimerkiksi tietokantakyselyillä tai liittimien (engl. data connector) avulla. ETL-prosessin aikana esimerkiksi datan epäsäännöllisyydet puhdistetaan, tyhjät arvot korvataan oletusarvoilla tai esimerkiksi keskiarvoilla, datalle suoritetaan muunnoksia ja sitä aggregoidaan. Tarvittaessa datan tietotyypit muunnetaan. Datasta poistetaan varmasti väärät tiedot ja kaksoiskappaleet. Tietokenttiä voidaan jättää tyhjiksi, jos arvon puuttumisen syy tunnistetaan tai jos puuttuva arvo täydentyy myöhemmin. Datan valmistelulla on suuri merkitys, sillä datan laatu vaikuttaa suoraan analyysin laatuun. Data on tarpeeksi laadukasta analyysiä ja tiedon tuonnin automatisointia varten, jos se on ajantasaista, virheetöntä, yhdenmukaista sekä johdonmukaista. ETL-prosessi on toistuva tapahtuma. (Loshin 2012, 186, 192, 194–195; Maheshwari 2014, 41, 48–49; Sherman 2014, 13–17.)

4 BI-TYÖKALUT

BI-työkalut ovat BI:n loppukäyttäjälle näkyvä osa. Niiden avulla voidaan analysoida reaaliaikaista tilannetta, havaita datasta esille nousevia kehityssuuntia sekä tarkastella suuria tietokokonaisuuksia tai syventyä yksityiskohtiin. BI-työkaluja käytetään analyysin lisäksi tiedon välittämiseen loppukäyttäjille eli raportointiin. Taulukkolaskentaohjelmat ovat yksinkertaisimpia BI-työkaluja, joilla voidaan analysoida dataa eri lähteistä. Myös taulukkolaskentaohjelmia voidaan automatisoida rajatusti esimerkiksi makrofunktioilla. Edistyneemmillä BI-työkaluilla laaditaan raportteja, tiedon koontinäyttöjä ja esitetään ne loppukäyttäjille graafisessa käyttöliittymässä. Samalla piilossa ovat tiedonhaku ja monimutkaiset tilastolliset laskennat. (Maheshwari 2014, 24–25; Sherman 2014, 13–18, 95–96, 150–151.)

4.1 Tietomallit BI-työkaluissa

Tietomalli on käsite, joka tulee usein esille BI-työkaluista puhuttaessa. Seuraavaksi kerrotaan BI-työkaluihin liittyvistä tietomalleista yleisemmin. Kappaleen lopussa kerrotaan tarkemmin dimensiomallin eri tapauksista.

Tietomallit ovat visuaalisia esityksiä järjestelmässä olevasta tiedosta ja niiden välisistä suhteista. Relaatiotietokannat kuten operatiivisten järjestelmien tietokannat hyödyntävät usein ER-mallia (engl. entity relationship), jotta tietokantakyselyt ovat nopeita ja tehokkaita. Normalisoidussa mallissa pyritään pitämään datan toistuvuus mahdollisimman vähäisenä, ja datan on oltava mahdollisimman ehyttä. (Hamsen & Garcia 2012, 114; Loshin 2012, 47; Sherman 2014, 107, 173–176, 178–179, 190.)

Tietovarastot ja BI-työkalut käyttävät puolestaan niin kutsuttua dimensiomallia. Dimensiomalli on suunniteltu tukemaan tiedonhakua, analysointia ja raportointia. Dimensiomalli on helpommin ymmärrettävä, sillä se kuvaa loogisemmin liiketoimintaa. Dimensiomallissa avaintiedot ovat keskeisessä taulussa, jota kutsutaan faktatauluksi. Jokainen faktataulun fakta on jokin havainnoitava ja mitattava asia. Faktataulua täydentävät lisätietotaulut, eli dimensiotaulut, jotka kuvailevat faktoja. Dimensioita käytetään datan ryhmittelyyn ja suodattamiseen. Tiedonhaku on nopeampaa, koska tietoa haetaan kootusti faktataulusta sen sijaan, että tietoa haettaisiin useista yhteen liitetyistä tauluista, joita normalisoidun tietorakenteen käyttäminen vaatii. Dimensiomallissa faktat ja dimensiot on liitetty toisiinsa avaimien avulla. Dimensiomallia voidaan laajentaa, kunhan faktatauluihin lisätään avain uudelle dimensiolle. Dimensiomallin faktataulut on normalisoitu, mutta dimensiot ovat denormalisoidussa muodossa. (Loshin 2012, 109–111; Maheshwari 2014, 38–39; Sherman 2014, 197–203, 205, 208–209.)

ER-mallista voidaan muodostaa dimensiomalli jakamalla ER-malli liiketoiminnallisiin prosesseihin. Jokainen prosessi mallinnetaan erikseen. Avaintekijät tunnistetaan mittareiksi ja näistä tehdään faktataulu. Jäljelle jääneistä tauluista tehdään dimensiotauluja. (Hamsen & García 2012, 117–118.)

Yleisesti ottaen dimensiomallissa on vähemmän entiteettejä kuin ER-mallissa, sillä raportointiin ja analytiikkaan ei tarvita kaikkia niitä tietoja, mitä operatiivisten järjestelmien toimintaan tarvitaan. Dimensiomallin voi muodostaa kolmella eri skeemalla. Skeemat ovat tähtimalli, lumihutalemalli ja moniulotteinen malli. Skeeman valinta riippuu siitä minkälaista analyysiä ja raportointia datalla tehdään, mitä rajoituksia sille on asetettu ja mikä BI-työkalu on käytössä. (Sherman 2014, 197–203, 208, 213.)

Tähtimalli muodostuu faktataulusta ja sen ympärillä olevista useista dimensiotauluista. Tähtimalli on saanut nimensä siitä, että se muistuttaa tähteä. Tähtimallissa faktataulu on normalisoidussa muodossa, jotta siihen suoritettavat kyselyt olisivat tehokkaita. Attribuutit voivat olla jaettuna useisiin eri dimensiotauluihin. Tähtimallissa voidaan ajatella, että dimensiotaulut suodattavat faktataulun tietoja. (Loshin 2012, 110; Maheshwari 2014, 42; Sherman 2014, 208–209; Arnold 2022, 65.)

Lumihutalemalli poikkeaa tähtimallista niin, että myös dimensiotauluilla on niitä tarkentavia dimensiotauluja. Lumihutalemallisakin faktatauluun liitetyt dimensiotaulut ovat denormalisoidussa muodossa, mutta dimensioita kuvaavat hierarkiat voivat olla normalisoidussa muodossa ja jokainen hierarkia muodostaa oman dimensiotaulunsa. Faktataululla on yhteys jokaiseen hierarkian dimensiotauluun korkeimmasta tasosta matalimpaan. Osa BI-työkaluista on suunniteltu hyödyntämään erityisesti lumihutalemallia. (Hamsen & García 2012, 117; Maheshwari 2014, 42; Sherman 2014, 201.)

Moniulotteinen malli on moniulotteinen taulukkorakenne, jossa taulukosta on haettavissa kaikki yksityiskohtaiset tiedot sekä tunnuslukujen aggregaatiot valmiina. Moniulotteisessa mallissa jokainen solu on mittari. Taulukon dimensiot ja hierarkiat on kuvattu moniulotteisen mallin metadatanä. Moniulotteisen mallin skeemat tallennetaan moniulotteisiin tietokantoihin, joita kutsutaan OLAP-kuutioiksi (engl. online analytical processing cube). Myös OLAP-kuutiolle voi tehdä dataan porautuvia kyselyitä. (Sherman 2014, 211.)

4.2 Datan visualisointi

Datan visualisointi on datan esittämistä graafisessa muodossa kuten kuvaajina, jotta suuren tietomäärän omaksuminen olisi helpompaa verrattuna taulukkomuotoiseen dataan. Kuvaajista voidaan helpommin tunnistaa poikkeavia tuloksia. Visualisointeja tehdessä tulisi valita analyysiin sopivat kuvaajat, sillä epäsopivat kuvaajat voivat aiheuttaa vääriä tulkintoja. Datan visualisoinnissa on huomioitava, että ihminen pystyy omaksumaan rajallisen määrän tietoa kerrallaan. Siksi kuvaajien ja taulukoiden tulisi nostaa esille informaation ydinasia ja johtopäätökset. Kuitenkin on huomattava, että visualisointien laatu on riippuvainen niihin käytettävän datan laadusta. (Loshin 2012, 72; Maheswari 2014, 15; Sherman 2014, 17, 401–402.)

Käytettävät visualisoinnit riippuvat siitä, minkälaisesta datasta on kyse. Yleisimmät analyysit ovat erilaisia vertailuja, aikasarjoja tai trendejä. Näiden lisäksi voidaan visualisoida muun muassa korrelaatioita, jakaumia tai maantieteellistä jakaumaa. Pylväs- ja palkkikuvaajat soveltuvat vertailuun. Ne ovat hyödyllisiä, jos verrataan arvoja keskenään eri kontekstissa. Suuruusjärjestyksen voi visualisoida järjestämällä palkit koon mukaan. Aikasarjojen vertailuun soveltuvat erilaiset viiva-, pinta-ala- ja pylväskuvaajat. Viivakuvaajista näkee arvojen muutoksen ajan suhteen. Kokonaisuutta voidaan kuvata esimerkiksi piirakkakuvaajalla, jossa yksi siivu vastaa osuutta kokonaisuudesta. Piirakkakuvaaja on soveltuva, jos data on jaettu muutamaan kategoriaan. Piirakkakuvaaja voidaan korvata esimerkiksi lämpökartalla, jos kategorioita on paljon. Hajontakuvaajalla voidaan havainnollistaa muuttujien välisiä yhteyksiä. Kuplakuvaaja on hajontakuvaajien erikoistapaus, johon on lisätty kolmas ulottuvuus kuten lukumäärä. Kartat ovat puolestaan hyvä tapa visualisoida dataa, johon liittyy sijaintitietoja. Mittarikuvaajalla (engl. gauge) voidaan seurata arvoa suhteessa raja-arvoihin. Hämähäkkikaavio esittää muuttujat eri ulottuvuuksien sarjassa. Näiden lisäksi voidaan käyttää myös pienoisiivakaaavioita, joissa data esitetään yksinkertaistetusti ilman akseleita tai koordinaatteja. Penoisiivakaaavioita voidaan käyttää trendien vertailuun. Suuren datamäärän tarkasteluun sopii parhaiten taulukko tai matriisi, ja erityisesti jos niihin hyödynnetään ehdollista muotoilua esimerkiksi värien avulla. Matriisissa voidaan esittää myös kategorioita ja hierarkioita. Puukartat ovat hyviä kuvaamaan hierarkioita. (Loshin 2012, 312–314; Sherman 2014, 351–355; Knight ym. 2018, 108–134.)

Visualisointeja valittaessa tulisi kiinnittää huomiota ruututilan rajallisuuteen ja sen riippuvuuteen käytettävästä laitteesta. Tämän vuoksi visualisointeja tulisi käyttää harkiten, eli valita kontekstiin sopiva visualisointi ja valita kuvaajia, jotka välittävät viestin tehokkaasti. Ylimääräinen visuaalinen kuorma ei tuota lisäarvoa BI:n kannalta, ja esitys olisi hyvä pitää mahdollisimman yksinkertaisena. (Loshin 2012, 316–317; Canziba 2018, 315–317.)

Värivalintoja tehdessä tulisi tiedostaa, että väreihin yhdistyy alitajuisesti eri tunnetiloja ja merkityksiä. Hyvä esimerkki on liikennevalojen värit. Vihreä väri voi esittää luontoa, energiaa ja muita niihin liittyviä asioita. Punainen väri voi esittää vaaraa ja vihan tunteita. Sininen puolestaan yhdistetään usein rauhallisuuteen, kun taas keltainen väri yhdistetään usein iloon tai lämpöön. Käyttäjät voivat tulkita värejä eri tavalla esimerkeistä poiketen. (Canziba 2018, 141.)

Graafisia elementtejä voidaan hyödyntää tiedon ryhmittelyssä ja järjestelyissä. Niiden avulla voidaan ohjata katsetta esimerkiksi käyttämällä viivoja ja nuolia. Ympyrät symbolisoivat päättymätöntä kiertoa, tai kokonaisuutta ja yhteneväisyyttä. Neliöitä voidaan hyödyntää yhtenäisyyden, turvallisuuden ja tasa-arvon kuvaamisessa. Spiraalit puolestaan voivat symboloida kasvua tai laajenemista. (Canziba 2018, 135–139.)

4.3 Power BI

Power BI on Microsoftin tuoteperhe, jonka avulla on mahdollista laatia analyysejä ja raportteja useista eri tietolähteistä ja jakaa niitä loppukäyttäjille. Power BI mahdollistaa tiimityöskentelyn yhteisissä työtiloissa, ja sen avulla voidaan määrittää tarkasti, kenellä on pääsy jaettavaan tietoon. Sen avulla analyytikot voivat tehdä raportteja, joista loppukäyttäjä pystyy suodattamaan haluamansa tiedon interaktiivisesti klikkaamalla visualisointeja. Power BI -tuoteperhe koostuu seuraavista osista: Power BI Desktop, Power BI -palvelu, Power BI Mobile, Power BI Report Builder, Power BI Report Server on Premises ja Power BI Embedded. Power BI -palvelu on saatavilla Azure-pilvipalvelussa ohjelmisto palveluna -mallin eli niin kutsutun SaaS-mallin mukaisesti (engl. software as service), Power BI Desktop on puolestaan Windows-työasemalle asennettava ohjelmisto. Power BI -mobiilisovellus on saatavilla Windows-,

iOS- ja Android-laitteille. (Knight ym. 2018, 7, 102; Arnold 2022, 2, 74, 181; Microsoft 2024b; Microsoft 2024e.)

Power BI Report Builder on käyttäjän tietokoneelle asennettava ilmainen sovellus sivutettujen raporttien laatimiseen. Power BI Embeddedillä voidaan upottaa Power BI -raportteja ja visualisointeja sovelluksiin. (Arnold 2022, 3; Microsoft 2024b.) Opinnäytetyössä tarkastellaan syvällisemmin Power BI Desktopia ja Power BI -palvelua.

4.3.1 Power BI Desktop

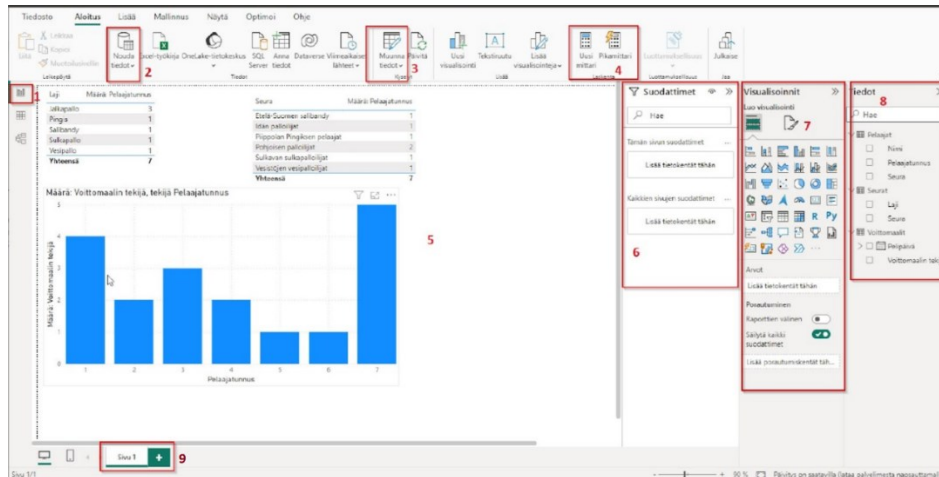
Power BI Desktop on raporttien kehittämiseen tarkoitettu työkalu, joka asennetaan paikallisesti käyttäjän tietokoneelle. Power BI Desktopissa tehdään datan visualisointi ja raportit, jotka jaetaan loppukäyttäjille Power BI -palvelussa. Power BI Desktopissa raportin tietojoukko rakennetaan tuomalla dataa eri tietolähteistä, minkä jälkeen määritetään, miten dataa käsitellään. Power BI Desktopissa laaditaan myös tietojoukon tietomalli yhteyksineen. Microsoft julkaisee Power BI Desktop -päivityksiä kuukausittain. (Knight ym. 2018 8–9; Arnold 2022, 3; Microsoft 2024b.)

Power BI Desktop käyttää kahta moottoria tiedon prosessointiin. Kaavamoottori käsittelee ja optimoi miten ja missä järjestyksessä data haetaan. Datan tallennukseen on oma moottori (VertiPaq), joka tallentaa ja hakee tietomallin datan, jota kyselymoottori pyytää. Toimintojen jakaminen omiin moottoreihin tehostaa ja nopeuttaa kyselyitä ja pienentää tiedostokokoa. (Arnold 2022, 9–10.)

Power BI Desktop rakentuu kahdesta osasta. Käyttöliittymään, jossa on raportit, tietojoukot ja tietomallit sekä Power Query -editoriin. Power Query -editoria käytetään datan hakemiseen ja muokkaamiseen. Raportit ja datan visualisointi tehdään käyttöliittymän klikkaa ja raahaa -toiminnoilla. (Arnold 2022, 4; Microsoft 2024b.)

Power BI Desktopin käyttöliittymä rakentuu seuraavista osista, jotka on esitetty kuvassa 1. Power BI:n raporttinäkymästä. Kuvassa 1 on näkyvillä raportin sisältö ja ulkoasu (Knight ym. 2018, 9; Microsoft 2024b).

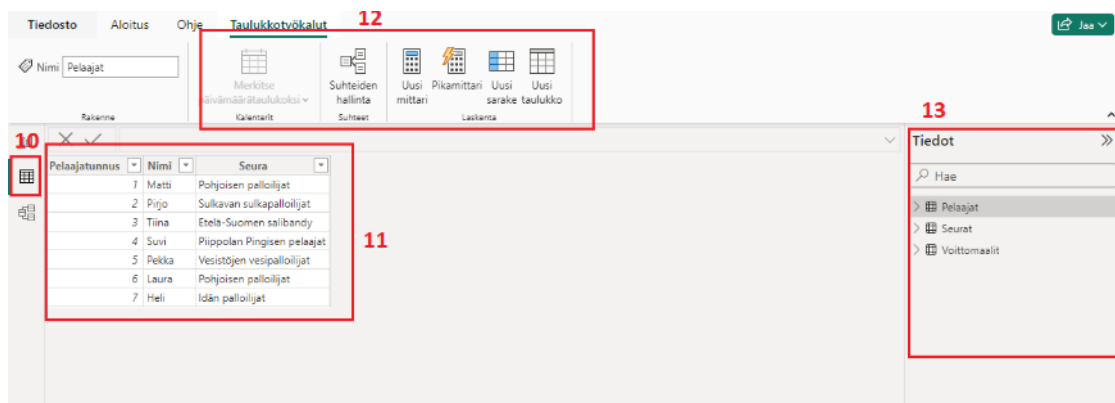
1. Raportinäkymän painike
2. Nouda tiedot -painike tietojen hakemista varten
3. Muunna tiedot -painike Power Query -editorin avaamista varten
4. Painikkeet uuden mittarin sekä pikamittarin luomista varten
5. Raporttinäkymä, jossa on esimerkin vuoksi palkkikuvaaja ja taulukoita
6. Suodatinvalikko
7. Visualisointivalikko
8. Tietovalikko
9. Raportin sivujen vaihto



Kuva 1. Power BI Desktopin raporttinäkymä

Kuvassa 2 on tietonäkymä, jossa on esitetty raporttiin käytettävä data taulukkomuodossa. Datasivulla on nähtävillä myös datasta tehdyt mittarit (Knight ym. 2018, 9; Microsoft 2024b).

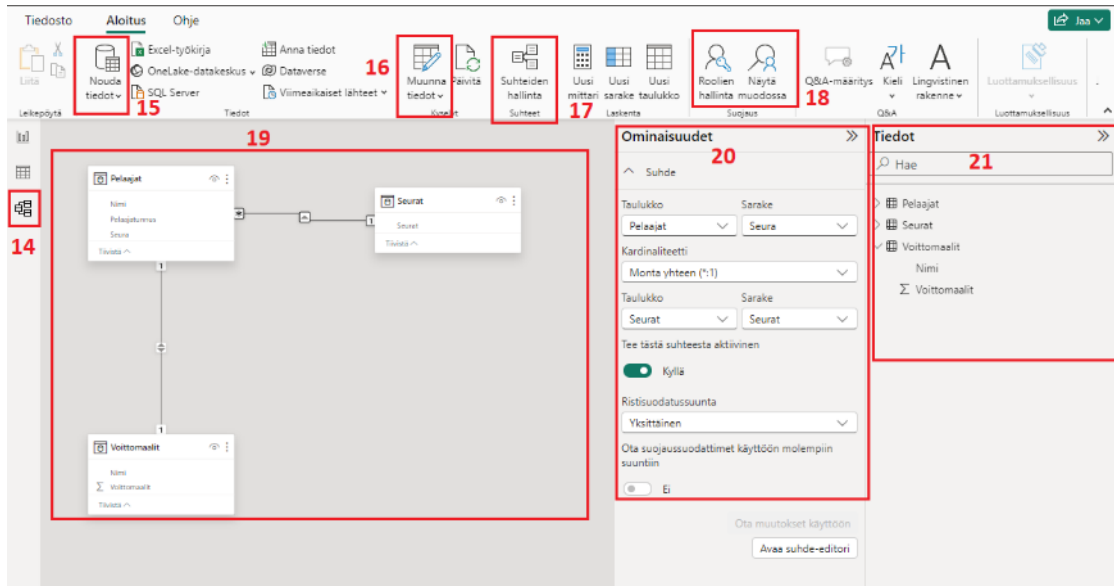
10. Taulukkonäkymän avauspainike
11. Raportin tiedot esitettynä taulukossa
12. Taulukkotyökaluja
13. Tietovalikko



Kuva 2. Power BI Desktopin tietonäkymä

Kuvassa 3 on tietomallinäköymä, jossa on esitetty tietojoukon tietomalli. Siellä voidaan luoda yhteydet tietotaulujen välille. Power BI:ssa yhteys muodostetaan kahden taulun sarakkeiden välille, ja sitä määrittää myös datan suodatukseen suunta sekä yhteyden kardinaliteetti. Tietomallin taulujen väliset yhteydet mahdollistavat Power BI:n automaattisen suodatuksen ja mittarit. Power BI havaitsee ja luo automaattisesti mahdollisia yhteyksiä, mutta ne tulisi aina tarkistaa. Taulujen välillä voi olla kerrallaan vain yksi aktiivinen yhteys, vaikka passiivisia yhteyksiä voi olla monta. (Knight ym. 2018, 9, 49; Arnold 2022, 29, 58; Microsoft 2024b.)

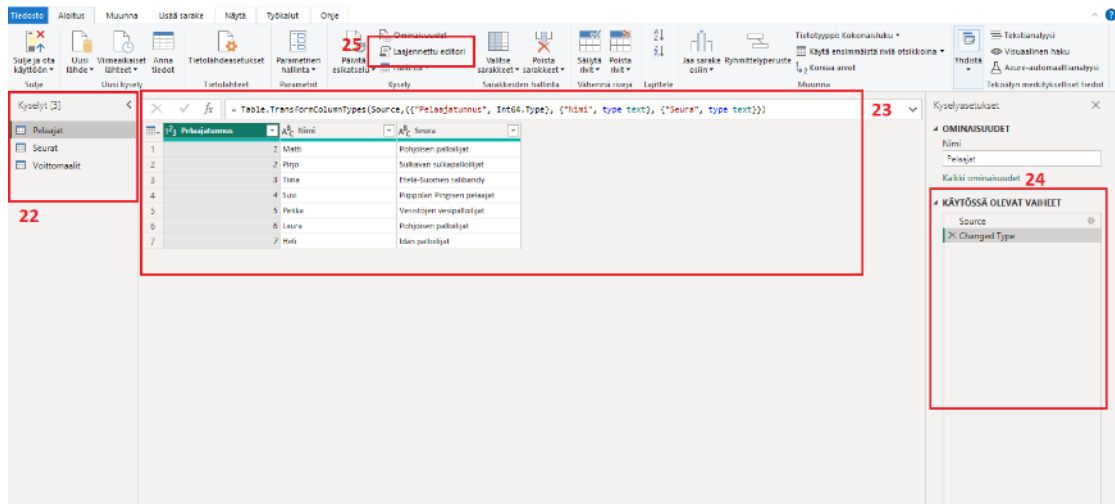
14. Tietomallin avaamispainike
15. Tietojen noutamisen painike
16. Tietojen muuntamisen painike
17. Tietomallin suhteiden hallinnan painike
18. Roolien hallintaan ja esikatseluun liittyvät painikkeet
19. Tietojoukon tietomalli
20. Tietomallin ominaisuudet -valikko
21. Tiedot-valikko



Kuva 3. Power BI Desktopin tietomallinäköymä

Kuvassa 4 on esitetty Power Query -editorin näköymä (Microsoft 2024g).

22. Tietojoukkoon käytettävät kyselyt
23. Tietojoukon esikatselu ja komentorivi
24. Valitussa kyselyssä käytössä olevat vaiheet
25. Laajennettu editori M-kielen kirjoittamista varten



Kuva 4. Power Query -editori

Kuvista 1–4 nähdään, että Power BI Desktopin käyttöliittymän rakenne pysyy samankaltaisena jokaisessa näkymässä. Työkalupalkki on käyttöliittymän yläosassa. Tietoa käsittelevät valikot on sijoitettu sivun oikeaan reunaan. Varsinainen työtila on käyttöliittymän keskellä.

Power Query, M ja muut kielet

Power BI:n data tuodaan, muunnetaan ja puhdistetaan analysointi varten Power Query -editorilla. Power Query helpottaa datan käsittelyä, sillä yksinkertaisissa tapauksissa Power Query -editorin käyttöliittymään rakennetut toiminnot ovat riittäviä datan muuntamiseen. Power Queryllä datan muunnokset kohdistuvat taulun sarakkeisiin eikä yksittäisiin soluihin. Power Queryn taustalla kirjoitetaan automaattisesti M-kieltä. M-koodia on mahdollista muokata tai kirjoittaa itse, jos valmiit toiminnot eivät ole riittäviä datan käsittelyyn. Kaikki datalle tehtävät muunnokset tallentuvat omiksi vaiheiksi Power Query -kyselyyn. Vaiheiden järjestystä voidaan muuttaa ja niitä voidaan poistaa tai lisätä. Power Query voidaan käynnistää datan noutamisen yhteydessä tai myöhemmin Power BI Desktopin työkaluvalikoista. (Knight ym. 2018, 20–24, 45; Arnold 2022, 4,8; Microsoft 2024b.)

Power BI:llä voidaan käyttää myös tilastollisesta analyysistä juontavaa R-kieltä, jota käytetään tilastollisessa laskennassa. R-kielen funktioita ja kirjastoja voidaan käyttää Power Query -editorilla. R-kielen lisäksi Power BI voidaan

konfiguroida käyttämään myös Pythonia, joka on skriptikieli. R ja Python täytyy asentaa ja konfiguroida Power BI:lle erikseen. R-kieltä ja Pythonia käytettäessä on rajoitteena se, että data tulee päivittää henkilökohtaista dataväylää käyttäen paikalliselta koneelta. Power BI -palvelussa ei ole käytettävissä kaikkia R-kirjastoja ja Python-paketteja, joita käyttäjällä voi olla asennettuna omalle tietokoneelle. Datan tietoturva ajatellen on huomioitava, että tällöin tietolähteet tulee merkitä julkisiksi. (Knight ym. 2018, 42; Arnold 2022, 177–178.)

DAX

Power BI:n komentokieli DAX (engl. data analysis expressions) lanseerattiin vuonna 2010 osana Excel-taulukkolaskentaohjelmiston Power Pivot -lisäosaa. DAX-kieltä käytetään Power BI:n ja Power Pivotin lisäksi myös Analysis Services Tabular -ohjelmassa. DAX-kieli on samankaltaista kuin Excel-taulukkolaskentaohjelmiston laskukaavoihin käytettävä komentokieli, joten se on useimmiten helppo omaksua, jos käyttäjä tuntee entuudestaan Excelin. Power BI Desktopiin on sisäänrakennettu IntelliSense-toiminto, joka täydentää ja ehdottaa DAX-komentoja käyttäjän syötteen perusteella. DAXia käytetään usein mittareiden ja laskettujen sarakkeiden tekemiseen. Kaikkien Power BI:n visualisointien taustalla toimii myös DAX-kieli, jota Power BI kirjoittaa piilossa käyttäjältä. (Knight ym. 2018, 79; Arnold 2022, 10, 101–102.)

Power BI:n tietomalliin perustuva automaattista suodatustoimintoa voidaan muokata tarpeen mukaan DAX-kielellä käyttäen esimerkiksi CALCULATE-funktiota. Esimerkki DAX-komentokielen tehokkaista analyysitoiminnoista ovat kattavat aikatietofunktiot, joiden avulla voidaan tehdä erilaisia ajanjaksoja käsitteleviä analyysejä. (Knight ym. 2018, 92, 94; Microsoft 2023b.)

DAX-funktiot voidaan luokitella eri ryhmiin, joita ovat aggregaattifunktiot, päivämäärä- ja aikafunktiot, suodatusfunktiot, taloustietofunktiot, informatiiviset funktiot, loogiset funktiot, matemaattiset ja trigonometriset funktiot, yhteysfunktiot, tilastolliset funktiot, taulukkofunktiot, tekstifunktiot, aikatietofunktiot ja muut funktiot. Aggregaattifunktiot tekevät datasta yhteenvedon kuten summan, keskiarvon, minimin, maksimin, lukumäärän ja niin edelleen. (Arnold 2022, 103–104.)

Tiedon noutaminen

Power BI tukee yli 80 erilaista dataliitintä, joiden avulla dataa voidaan noutaa erilaisista tietolähteistä, joihin lukeutuvat useat tiedostotyypit, pilvipalvelut, internetsivut, tietokantamoottorit ja muut SaaS-palvelut. Power BI -raportti voi hyödyntää dataa kolmella eri tapaa, jotka ovat datan tuominen, DirectQuery ja LiveConnection. (Knight ym. 2018, 10.)

Datan tuonti on Power BI:n oletusasetus, missä Power BI hakee ja tallentaa datan rivit tietolähteestä Power BI:n sisäiseen muistivarastoon. Tämän teknologian nimi on xVelocity (myös nimellä VertiPaq), ja se käyttää paikallista käyttäjältä piilotettua SQL Server Analysis Services -virtuaalikonetta. Power BI:n xVelocity parantaa kyselyiden tehokkuutta, ja se pakkaa haetun datan pienempään tilaan. Datan tuonnissa tulee ottaa huomioon, miten usein ja millä tavalla dataa päivitetään, sillä data on staattista päivittämiseen saakka. Datan päivitys tehdään manuaalisesti, tai se voidaan ajastaa Power BI -palvelussa. Datan tuonnissa rajoitteeksi voi muodostua lähes reaaliaikaisen raportoinnin tarve. Datan tuonnissa tulee ottaa huomioon myös se, kuinka paljon dataa voidaan tallentaa sisäiseen muistiin. (Knight ym. 2018, 10–12; Arnold 2022, 9–10.)

DirectQueryn avulla Power BI yhdistyy suoraan tietolähteeseen. Tällöin dataa ei tallenneta Power BI Desktopin muistiin. DirectQuery voi olla hyvä valinta, jos tiedon tuonti aiheuttaa haasteita sisäisen muistin kanssa, käytettävä tietolähde on hyvin suuri tai analyysituloksia tarvitaan reaaliaikaisesti. Direct Queryn kanssa käyttäjän tietokoneen tehokkuus ei aiheuta samanlaisia haasteita kuin datan tuonnissa, sillä laskennat suoritetaan palvelimella. Analyysit ovat ajantasaisia, sillä tiedot haetaan reaaliaikaisesti. DirectQueryä ei kuitenkaan voi käyttää kaikkien tietolähteiden kanssa ja DirectQueryn tehokkuus riippuu käytettävästä tietolähteestä. Osa Power BI:n toiminnallisuuksista ei ole käytettävissä DirectQueryn kanssa, kuten jotkin Power Queryn toiminnallisuuksista. Osa DAX-funktioista eivät ole käytettävissä, erityisesti aikatiето-funktiot. Tämä johtuu siitä, että DAX:ia käytettäessä DirectQuery yrittää kääntää DAX-komennot tietolähteen natiiville kielelle. Estetyt mittareille sopivat DAX-funktiot on kuitenkin mahdollista erikseen sallia, vaikkakin tämä voi johtaa kyselyiden hidastumiseen. (Knight ym. 2018, 16–18.)

Kolmas tapa noutaa dataa Power BI:hin on Live Connection. Live Connection ei varsinaisesti nouda dataa Power BI:hin, vaan tällöin Power BI Desktopia käytetään datan visualisointityökaluna, joka kytketään olemassa olevaan tietojoukkoon. Live Connection ei kärsi tehokkuusongelmista, mutta rajoitteena on se, että se on tuettu vain muutamalla tietolähdetyypillä. Tuettuja tietolähdetyyppejä ovat SQL Server Analysis Services Tabular, SQL Server Analysis Services Multidimensional ja Power BI Service. Nämä käyttävät joko samaa xVelocity-moottoria kuin Power BI tai samankaltaista teknologiaa. Live Connectionia käytettäessä data päivittyy noin tunnin välein tai kun datan kanssa ollaan vuorovaikutuksessa. Päivitystiheyttä voidaan myös säätää. Live Connectionin rajoitteena on se, että tiedon mallintamiseen ja Power Queryyn liittyvät työkalut eivät ole käytettävissä, sillä nämä vaiheet on tehtävä aikaisemmin. (Knight ym. 2018, 19–20.)

4.3.2 Power BI -palvelu ja Power BI Report Server

Power BI -palvelu on Microsoftin Azure-pilvipalvelussa sijaitseva osa Power BI -kokonaisuutta. Power BI -palvelu itsessään on käytettävissä ilmaiseksi. Power BI -palveluun tunnistaudutaan Microsoft-tunnusten tai organisaation myöntämien käyttäjätunnusten avulla. Palvelu koostuu kahdesta osasta. Selainpuoli tunnistaa käyttäjän ja hallitsee tämän sisäänkirjautumisen palveluun. Palvelinpuoli käsittelee käyttäjän vuorovaikutuksen käyttöliittymän kanssa. Power BI käyttää Microsoft Entra ID:tä (entinen Azure Active Directory) käyttäjien hallinnoimisessa. (Arnold 2022, 3–5, 181; Microsoft 2024b; Microsoft 2024c.)

Power BI -palvelussa on oma raporttieditori, joka muistuttaa Power BI Desktopin käyttöliittymää. Power BI -palvelussa on mahdollista laatia raportteja käyttäjälle jaetuista tietojoukoista. Tietojoukkoja ja tietomalleja ei kuitenkaan voida muokata palvelussa toisin kuin Power BI Desktopin avulla. Tietojoukko voidaan julkaista omaan työtilaan tai esimerkiksi koko organisaation käyttöön. (Knight ym. 2018, 154–156; Arnold 2022, 181, 185; Microsoft 2024b.)

Jos organisaation tietoturvamääräykset eivät salli pilvipalvelun käyttöä, voidaan käyttää Power BI -palvelulle vaihtoehtoista organisaation omaan palvelinympäristöön asennettavaa Power BI Report Serveriä. Power BI Report Ser-

verillä jaeltavien raporttien kehitykseen tarvitaan oma Power BI Desktop -versio. Power BI Report Server päivitetään vain kolme kertaa vuodessa, mikä on vähemmän verrattuna Power BI Desktopin päivitystiheyteen. (Knight ym. 2018, 7, 171; Arnold 2022, 3; Microsoft 2024b.)

Työtilat

Raportit julkaistaan Power BI -palvelun työtiloihin (engl. workspace). Jokaisella käyttäjällä on oma yksityinen ja ilmainen työtila. Palvelussa on myös tiimityöskentelyyn tarkoitettuja työtiloja. Työtilaan voidaan julkaista useita raportteja joko Power BI Desktopin kautta, tai laatimalla raportti Power BI -palvelussa. Työtilassa voi myös tehdä koontinäyttöjä raporteista. Työtilan näkyyvyyttä voidaan hallita ja jakaa. Työtilarooleja ovat järjestelmänvalvojat, jäsenet, osallistujat sekä katselijat. Uuden työtilan luomiseen vaaditaan Power BI Pro -lisenssi. On suositeltavaa jakaa raportteja tiimityöskentelyyn tarkoitetun työtilan kautta oman työtilan asemesta. (Knight ym. 2018, 153–155, 163; Arnold 2022, 4–5, 181, 197–198; Microsoft 2024a; Microsoft 2024b; Microsoft 2024f.)

Työtiloille on saatavilla käyttöönottoputki. Käyttöönottoputki mahdollistaa raporttien kehittämisen ja testaamisen ennen julkaisua tuotantoympäristöön. Käyttöönottoputki on käytettävissä vain Premium-kapasiteetin työtiloissa. (Arnold 2022, 5, 195.)

Tiedon jakaminen ja suojaus

Power BI -palvelussa voidaan jakaa tietoa yksittäisillä raporteilla tai yhdistettynä koontinäyttöihin (engl. dashboard). Klikkaamalla koontinäytön visualisointia loppukäyttäjä pääsee tarkastelemaan yksityiskohtaisempaa raporttia. Koontinäyttöjen avulla voidaan jakaa myös reaaliaikaista dataa esimerkiksi Azure Stream Analytics -palvelun kautta. Jakamiseen vaikuttaa kuitenkin käyttäjien ja työtilojen lisenssit. (Knight ym. 2018, 157–159, 162–163.)

Tiedon jakeluun soveltuu Power BI -sovellus, kun käyttäjille halutaan jakaa raportteja eikä koko työtilaa. Sovelluksen avulla voidaan hallita raporttien näky-

vyttä eri yleisöille. Power BI -sovellukseen voidaan halutessa valita osa työtilan kaikista raporteista. Tämän lisäksi sovellukseen voidaan määrittää eri käyttäjäoikeudet kuin itse työtilaan. Tietojoukkoa ei jaeta sovelluksen mukana, vaikka raportit ja koontinäytöt jaetaan. Tietojoukko on tällöin piilossa muilta kuin työtilan jäseniltä. (Knight ym. 2018, 163; Arnold 2022, 193, 205; Microsoft 2024b.)

Power BI -palvelussa on mahdollista rajata tietojoukon ja raportin tietojen näkyvyyttä rivikohtaisesti käyttämällä rivikohtaista suojausta eli RLS:ää (engl. row level security). Power BI Desktopissa luodaan roolit, joille halutut rivit näkyvät. Näkyvyys määritetään DAX-kielellä. Power BI -palvelussa ihmisiä tai ryhmiä voidaan nimetä rooleihin. On huomionarvoista, että RLS ei toimi, jos työtilan jäsenellä on muokkausoikeus tietojoukkoon. Objektien suojauksella eli OLS:llä (engl. object level security) voidaan suojata jotkin taulut tai sarakkeet katselijoilta. Opinnäytetyön kirjoitushetkellä OLS:ia ei voi määrittää Power BI Desktopissa ja nämä roolit täytyy tehdä ulkoisilla työkaluilla kuten Tabular Editorilla. (Knight ym. 2018, 164–166; Microsoft 2024c.)

Power BI:ssä tietojoukot, raportit ja koontinäytöt voidaan merkitä luottamuksellisuustunnisteilla, jotka ovat käytössä Microsoftin Purview Information Protection -ohjelmistossa. Luokittelu siirtyy Power BI -datasta myös tiedon viennin yhteydessä Excel-, Power Point- ja PDF-tiedostoihin sekä muihin yhteensopiiviin tiedostomuotoihin. Luottamuksellisuustunnisteet voidaan määrittää Power BI Desktopissa ja Power BI -palvelussa. Luottamuksellisuustunnisteet eivät vaikuta pääsyyn itse dataan. Dataan pääsy määritetään Power BI:n pääsoikeuksilla dataan ja työtiloihin. Luottamuksellisen datan monitorointi on mahdollista pilvipalvelussa Microsoft Defenderin avulla. (Microsoft 2024c.)

Tietolähteiden yksityisyystaso kertoo onko data julkista, yksityistä tai organisaation sisäistä dataa. Yksityystaso määrittää miten data eristetään toisista datalähteistä. Esimerkiksi organisaation tietoa voidaan yhdistää yksityiseen dataan, mutta yksityistä dataa ei voida yhdistää organisaation dataa. Yksityistä dataa on esimerkiksi henkilötiedot, organisaatitietoa on organisaation sisäinen data, julkiset tietolähteet ovat esimerkiksi internetissä saatavilla olevaa dataa. (Microsoft 2024c.)

Saavutettavuus

Power BI:ssä on sisäänrakennettuja saavutettavuustoimintoja, joista osa vaatii konfigurointia. Ilman konfigurointia käytettävissä ovat näytönlukijoiden yhteensopivuus Power BI:n kanssa, näppäimistö navigointi sekä korkeakontrastinen värinäkö, keskittymistila ja taulukkonäkymät. Erillistä konfigurointia vaativat vaihtoehtoisen tekstin käyttäminen, välilehtien järjesteleminen ja erilaisten otsikkojen ja tunnisteiden käyttäminen. (Microsoft 2023a.)

4.3.3 Lisensointi

Power BI -palvelu toimii niin sanotun Freemium-mallin mukaisesti, eli siinä käyttäjät saavat lähestulkoon kaikki palvelun toiminnot ilmaiseksi. Power BI -lisenssimalliin kuuluvat ilmainen lisenssi, Pro-lisenssi, Premium per User, ja Premium Per Capacity -lisenssit. Omien raporttien ja tietojoukkojen jakamiseen sekä tiimityöskentelyyn tarvitaan vähintään Power BI Pro -lisenssi, sillä raportti voidaan jakaa suoraan käyttäjälle, jolla on vähintään Power BI Pro -lisenssi. Työtilaan voidaan kutsua vähintään Power BI Pro -lisenssin omaava käyttäjä. Power BI Pro -lisenssi sallii suurempien tietojoukkojen lataamisen palveluun sekä tietojoukkojen päivittämisen useammin kuin ilmainen lisenssi. (Knight ym. 2018, 154–156; Arnold 2022, 4–5, 181, 209–211; Microsoft 2024b.)

Power BI Pro -lisenssi tuo mukanaan joitain rajoituksia. Tiedostokoko voi enimmillään olla 1 GB. Tietojoukkoa voi päivittää maksimissaan 8 kertaa päivässä. Power BI Pro -lisenssillä integrointi Azuren tekoälytyökaluihin ei ole käytettävissä. Työtilojen osalta käyttöönottoputket eivät ole saatavilla ja niiden yhteiskoko ei saa ylittää 10 GB kaikkien työtilojen osalta. Sivutetut raportit eivät ole Pro-lisenssillä käytettävissä. (Arnold 2022, 210–211.)

Premium Per User -lisenssi on Pro-lisenssin ja Premium per Capacity -lisenssien välimaastossa. Se voi tulla tarpeeseen, jos Pro-lisenssin kattamat mahdollisuudet eivät riitä esimerkiksi liian suuren tietojoukon takia tai tietojoukkoa tulisi päivittää tiheämmin. Premium Per User -lisenssillä on käytettävissä myös Azuren tekoälypalveluita ja käyttöönottoputki. On kuitenkin huomioitava, että Premium per User -käyttäjän tekemästä työtilasta jaetut raportit voi jakaa vain

Premium per User -lisenssin omaaville käyttäjille. (Arnold 2022, 211–212; Microsoft 2024a; Microsoft 2024b.)

Premium Per Capacity -lisenssi sisältää kaikki Premium Per User -lisenssin toiminnot, mutta sen lisäksi organisaatiolle on varattu muistia sekä laskentaytimien laskentatehoa Power BI -palvelusta. Premium per Capacity on saatavilla tasoille P1-P5. Kapasiteetti on asiakkaalle varattu laskentaresurssi. Esimerkiksi P1/A4 -tason kapasiteettiin kuuluu 8 v-ydintä, tietojoukon maksimikoko voi olla 25 GB. Työtilat ovat oletuksena jaettujen kapasiteettien sisällä, mutta niin omia kuin muitakin työtiloja voidaan osoittaa Premium-kapasiteetin alle. Premium per Capacity -työtilat voidaan lisätä osaksi käyttöönottoputkea. Premium Capacity tai Premium per User -käyttäjäoikeudella on saatavilla Data-martit. Sivutetut raportit ovat saatavilla vain Premium per User tai Premium Per Capacity -työtilojen lisensseillä. (Arnold 2022, 187–188, 195, 201–212; Microsoft 2024b; Microsoft 2024c.)

Opinnäytetyön kirjoitushetkellä käyttäjäkohtainen Power BI Pro -lisenssi maksaa 9,40 €/kk. Käyttäjäkohtaisen Power BI Premium -lisenssin hinta on 18,70 €. Power BI Premium kapasiteettia kohti maksaa 4675,60 €/kapasiteetti/kuukausi varastoyksikölle P1. (Microsoft 2024d.)

4.4 QlikView

QlikView'n on kehittänyt ruotsalainen yhtiö QlikTech. QlikView'tä käytetään tiedon hakuun, käsittelyyn, analysointiin sekä jakeluun liiketoiminnan analytiikoille ja loppukäyttäjille. QlikView'llä dataa voidaan tutkia joustavasti käyttäjän kiinnostuksen mukaan, sillä käyttäjien tekemät valinnat, kuten kuvaajien ja taulukoiden klikkaaminen tai valintalaatikoiden käyttö suodattavat kaikkea tietomallin dataa. Datan analysointi perustuu QlikView'n assosiatiiviseen moottoriin, joka luo datasta uusia näkymiä. (Hamsen & García 2012, 23–26, 33–36; QlikTech International AB 2023.)

Opinnäytetyön kirjoitushetkellä QlikTechin tuote Qlik Sense korvaa QlikView'tä, joka on ensimmäisen sukupolven BI-työkalu. QlikView'tä ei voida hyödyntää data-analysoinnin itsepalveluratkaisuna, eikä sitä voi käyttää datan louhintaan, eikä data valmisteleminen visuaalisesti ole mahdollista.

QlikView’ssä ei ole tekoälyä hyödyntäviä toimintoja, eikä sillä voida prosessoida luonnollista kieltä. Nämä ominaisuudet ovat saatavilla Qlik Sense -työkalussa. (QlikTech International AB 2023.) Opinnäytetyössä käsitellään kuitenkin QlikView-työkalua, ja Qlik Sense -työkalua ei esitellä enempää.

4.4.1 QlikView-tuoteperhe ja arkkitehtuuri

QlikView-tuoteperheeseen kuuluu erilaisia työkaluja sisällön luomiseen, päivittämiseen, julkaisemiseen ja jakeluun. QlikView-sovelluksia tai niin kutsuttuja dokumentteja tehdään QlikView Desktop -kehitystyökalulla. Dokumentteja voidaan päivittää QlikView Publisherin ja QlikView Serverin kautta. Dokumentteja voidaan jakaa loppukäyttäjälle QlikView Accesspointilla, joka on käytettävissä yleisimmillä verkkoselaimilla. Dokumentteja voidaan tarkastella myös QlikView Serverillä sekä QlikView Desktopilla. (Hamsen & García 2012, 43–45; QlikTech International AB 2023.)

QlikView Desktop on Windows-käyttöjärjestelmälle asennettava ohjelmisto, jota käytetään QlikView-dokumenttien kehittämiseen ja niiden tarkasteluun. QlikView Desktop vaatii vähintään 4 GB muistia, mutta tarvittava muisti riippuu suoraan datan määrästä. QlikView Desktop on saatavilla myös henkilökohtaisena versiona. (Hamsen & García 2012, 44–45; QlikTech International AB 2023.)

QlikView Server on palvelin, jonne voi keskitetysti sijoittaa QlikView-dokumentteja. Käyttäjät voivat avata QlikView Serverille ladattuja dokumentteja esimerkiksi verkkoselaimen kautta. QlikView Server tekee dokumenttiin liittyvät laskennat ja datan käsittelyn, minkä vuoksi käyttäjäkohtaista suurta kone-tehoa tai RAM-muistia ei tarvita. (Hamsen & García 2012, 44–45; QlikTech International AB 2023.)

QlikView Publisher on lisensoitu moduuli QlikView Serverillä. Sitä käytetään datan lataamiseen dokumentteihin. QlikView Publisherin avulla voidaan aika-tauluttaa datan päivitys. Sama voidaan tehdä myös QlikView Serverin avulla. QlikView Publisherilla hallitaan sisältöä ja käyttäjien pääsyä dokumentteihin. QlikView Publisher on integroitu QlikView Serveriin ja QlikView Management -

konsoliin. QlikView Workbench on toinen QlikView Serverin lisensoitu moduuli. Se vaatii tiedonhallintajärjestelmän kuten Microsoft SharePoint 2019 sekä kehittäjätyökalun kuten Microsoft Visual Studio 2019. (Hamsen & García 2012, 44–45; QlikTech International AB 2023.)

4.4.2 QlikView-dokumenttien kehittäminen

QlikView-dokumentit tai QlikView-sovellukset sisältävät datan, tehdyt analyysit ja raportit. Dokumentissa on myös skripti, eli komennot, joita QlikView käyttää. Skriptien avulla haetaan data tietolähteistä ja käsitellään sitä halutulla tavalla. QlikView tarjoaa erilaisia avustajatoimintoja (engl. wizard) dokumenttien tekemiseen. QlikView-dokumentit tallennetaan QlikView'n qvw-tiedostomuotoon. Dokumentteja voidaan versioda muun muassa QlikView Project -tiedostoilla, joihin tallentuu sovelluksen muutokset, mutta ei kuitenkaan itse data. (Hamsen & García 2012, 52–53; QlikTech International AB 2023.)

QlikView-dokumentissa datan käsittely tapahtuu taustalla. Dokumentissa on myös käyttöliittymä, jota loppukäyttäjä käyttää datan tarkasteluun ja analysointiin. Käyttöliittymässä on mukana kaikki dokumentin objektit kuten kuvaajat ja valintalistat. Dokumentit voidaan organisoida esimerkiksi välilehtiä hyödyntäen. (Hamsen & García 2012, 32, 54.)

Tiedon tuonti QlikView-dokumenttiin ja yhteensopivat tietolähteet

QlikView on yhteensopiva monen tietolähteen kanssa. Tietokannoista yleisimmät tietolähteet ovat ODBC- (engl. open database connectivity) ja OLE DB-tietokantoja (engl. object linking and embedding database). Tietokantayhteyttä varten tarvitaan yhteensopivat ajurit. Ajurit asennetaan kyselyt suorittavalle koneelle tai palvelimelle. QlikView hakee datan tietokannasta SQL-kyselyllä. QlikView helpottaa kyselyn tekemistä Select Statement -avustajatoiminnolla, mutta komennon voi myös kirjoittaa itse. (Hamsen & García 2012, 27, 41–42, 56, 58, 80–81, 91–93.)

QlikView-dokumenttiin voidaan tuoda dataa eri tiedostomuodoissa kuten Excel-tiedostot, CSV-, TXT-, XML-, tai HTML-tiedostot ynnä muita tiedostomuotoja kuten QlikView'n omat QVD- ja QVX-tiedostot. Yhteensopivan tiedoston

tulee olla luettavissa ja sen rakenteen tulee olla selkeä. QlikView yrittää automaattisesti tulkita datan tiedostomuodon tiedostoavustajan avulla. QlikView voidaan yhdistää erilaisten liittimien avulla esimerkiksi SAPIin taikka sosiaalisen median sivustoihin. Liittimiä ja ajureita voidaan ohjelmoida itse tai ostaa palveluntarjoajalta. QlikTech tarjoaa liittimien ohjelmointia varten SDK:n (engl. software development kit). Osan liittimistä QlikTech tarjoaa lisenssimaksua vastaan. Ulkoisten tietolähteiden lisäksi dataa voidaan lisätä QlikView-dokumenttiin kirjoittamalla suoraan skriptiin Edit Script -ikkunan kautta. (Hamsen & García 2012, 27, 41–42, 56, 58, 80–81, 83, 85, 98–100, 110; QlikTech International AB 2023.)

Data tuodaan QlikView-dokumenttiin latausskriptillä. Latausskriptiin voidaan yhdistää dataa useista tietolähteistä. QlikView-dokumenttiin ladattu data tallennetaan RAM-muistiin tai se tallennetaan QVD-tiedostoksi. Ladattavan datan määrä on riippuvainen lähinnä käytettävän tietokoneen muistista. QVD-tiedosto on QlikView’lle natiivi ja optimoitu datatiedostomuoto. Sen prosessointi QlikView’illä on nopeaa ja tehokasta. QlikView pienentää datan viemää tilaa myös poistamalla datasta kaksoiskappaleet. (Hamsen & García 2012, 27, 41–42, 56, 58, 80–81; QlikTech International AB 2023.)

QVD-tiedoston käyttäminen pienentää datan viemää tilaa jopa 90 %. QlikView lataa datan QVD-tiedostosta noin 10–100 kertaa nopeammin kuin suoraan tietokannasta. QVD-tiedoston dataa voidaan käyttää myös useammassa QlikView-dokumentissa. (Hamsen & García 2012, 109; QlikTech International AB 2023.)

QVX-tiedosto eli QlikView eXchange -tiedosto on tiedostomuoto, johon tallennetaan ulkoisten tietolähteiden dataa. QVX on julkinen tiedostomuoto, joka muodostetaan ulkoisista tietolähteistä liittimien avulla. Se ei ole yhtä optimoitu kuin QVD-tiedosto. (Hamsen & García 2012, 109.)

QlikView’ssä dataa muokataan QlikView’n sisäänrakennetulla skriptikielellä. Datat muunnokset voidaan määrittää QlikView’n skriptieditorissa. Skriptissä komennot esitetään eri värillä tai fontilla riippuen komentosanojen merkityksestä. Skriptissä olevat yksittäiset lausekkeet loppuvat puolipisteeseen. Skripti

suoritetaan ylhäältä alas ja vasemmalta oikealle. Yksinkertaista datan muokkaamista varten voidaan käyttää apuna QlikView'n avustajatoimintoja. (Hamsen & García 2012, 246–248, 250, 336–337.)

Tietomalli

QlikView'n ydin on tietomalli, joka sisältää kaikki tietolähteen taulut. QlikView'n tietomallissa on assosiatiivinen yhteys kaikkien taulujen välillä, mikä mahdollistaa tietojen suodattamisen käyttäjän haluamalla tavalla. Assosiaatio on täysi ulkoliitos taulujen välillä. (Hamsen & García 2012, 55; QlikTech International AB 2023.)

Tietomallia voidaan tarkastella Table Viewer -ikkunassa. Table Viewer -ikkunassa tietomallin taulut on esitetty laatikkoina, joissa on taulujen sisältämät kentät. Taulujen välinen yhteys on merkitty viivalla laatikoiden välissä. Assosiaatio muodostuu automaattisesti kahden taulun välille, jos niillä on samanniminen kenttä. Tämän vuoksi QlikView'n luomat assosiaatiot tulee tarkistaa. Yhteys voidaan purkaa tai uusia yhteyksiä voidaan tehdä nimeämällä kenttiä uudelleen. Kahden taulun välillä tulisi olla vain yksi yhdistävä kenttä, sillä jos kenttiä on useampi luo QlikView synteettisen avaimen. Synteettiset avaimet voivat huonontaa suorituskykyä. (Hamsen & García 2012, 63–65, 97–98, 120–125; QlikTech International AB 2023.)

Koska QlikView poistaa datasta duplikaatit, ei sen tietomalliksi suositella lumihiutalemallia. Tähtimalli on QlikView'lle suositeltu tietomalli. QlikView on myös tehokkaampi ja nopeampi, kun taulujen välillä on vähemmän yhteyksiä (Hamsen & García 2012, 117–119.)

Tiedon esittäminen QlikView'ssä

QlikView-dokumenttiin voidaan lisätä paljon erilaisia objekteja datan esittämistä varten. Kaikkien objektien muokkaaminen tehdään dialogien avulla. Esimerkiksi tiedon suodattamiseen voidaan käyttää valintalistaa, joka sisältää oletuksena kaikki arvot mitä tietomallissa on. Valintalistalla voidaan tehdä valintoja ja suodatuksia. Kun dokumentissa on tehty valinta, päivittyvät kuvaajat

ja muut objektit sen mukaisesti. Käyttäjän tekemät valinnat säilyvät myös silloin, jos käyttäjä vie datan Excel-tiedostoon. Dokumentin valintoja tai näkymiä voidaan tallentaa myöhempää käyttöä varten QlikView'n kirjanmerkkitoiminnolla. (Hamsen & García 2012, 39, 58–59, 61, 67, 163.)

Kuvaajia lisätään dokumenttiin Create Chart -dialogilla. Dialogilla valitaan kuvaajan tyyppi ja siihen lisätään halutut dimensiot. Kuvaajalle lisätään myös laskentakaava, joka voidaan kirjoittaa tai valita pudotusvalikosta. (Hamsen & García 2012, 68–70, 77, 179–180.)

QlikView-dokumentissa voidaan ryhmitellä dataa eri tavoilla. Sykliset ryhmät kokoavat yhteen eri dimensioita. Porautuvia ryhmiä käytetään datan hierarkian määrittämiseen. Molemmat ryhmittelyt tehdään ryhmädialogien avulla. Raportissa nämä ryhmittelyt näkyvät ympyränuolikuvakkeella ja ylöspäin osoittavalla nuolikuvakkeella, joita klikkaamalla tarkasteltava dimensio vaihtuu. Käyttäjä voi myös lajitella datan haluamaansa järjestykseen esimerkiksi klikkaamalla taulukon sarakkeen otsikkosolua. (Hamsen & García 2012, 37, 64, 73, 76, 78.)

Tiedon suojaus

QlikView'ssä käyttäjän pääsy dataan määritetään joko dokumentin skriptissä tai QlikView Publisherin kautta. QlikView käyttää Windows-käyttöjärjestelmän käyttäjäautentikaatiota käyttäjän tunnistamiseen tai erillistä käyttäjätunnusta ja salasanaa. QlikView voi tehdä myös yksinkertaisen käyttäjän tunnistamisen lisenssiavaimen avulla. Käyttäjän valtuutus dataan voidaan asettaa QlikView'n skriptissä tai Windows-käyttöjärjestelmän avulla. (QlikTech International AB 2023.)

QlikView-dokumentit täytyy erottaa omiksi käyttäjä- tai käyttäjäryhmäkohtaisiksi tiedostoiksi, jos käyttäjien pääsyä hallitaan QlikView Publisherin avulla. Tällöin dokumentit tallennetaan kansioihin, joihin määritetään käyttöjärjestelmän turvallisuusasetukset, ja käyttöjärjestelmä hoitaa käyttäjien tunnistamisen ja valtuutuksen. Pääsyräjoitukset eivät ole käytössä, jos käyttäjä saa ladattua dokumentin toiselle käyttöjärjestelmälle. (QlikTech International AB 2023.)

Kun turvallisuusmääritykset tehdään QlikView-dokumenttiin, voidaan esimerkiksi kirjautumistunnisteet piilottaa muilta kehittäjiltä niin kutsutulla piilotetulla skriptillä, joka on skriptin salasanasuojattu osa. Piilotettua skriptiä ei näytetä ja sen tuloksia ei tallenneta QlikView'n lokitiedostoon. (Hamsen & García 2012, 514.)

QlikView-dokumentissa käyttäjien pääsyä dataan voidaan rajata lohkoittain (engl. section access), jolloin pääsy tietolohkoihin eli datan käyttöoikeudet määritetään QlikView'n skriptissä. Skriptiin voidaan määrittää esimerkiksi mitä kenttiä tai tauluja käyttäjä saa nähdä. Käyttäjältä voidaan myös kieltää kenttiä. QlikView:ssä tulee tehdä kielletyn datan poisto datan latauksen aluksi, jolloin QlikView poistaa kaiken sen datan, johon käyttäjällä ei ole pääsyä. Yksi dokumentti voi sisältää dataa eri käyttäjille tai käyttäjäryhmille, jos pääsy dataan on määritetty lohkoittain skriptissä, mutta käyttäjä näkee vain hänelle sallitun datan. Jos käyttäjän valtuutus määritetään skriptissä, tulee se osa skriptiä erottaa datan latauksesta. (Hamsen & García 2012, 517, 522, 525; QlikTech International AB 2023.)

Datan lisäksi voidaan määrittää dokumenttitasoinen tai välilehtitasoiset luvat käyttäjille. Näihin määrittäisiin kuuluu mitä käyttäjä voi dokumentilla tehdä kuten kuvaajien lisääminen tai poistaminen. Esimerkiksi ylläpitäjä voi muuttaa kaikkea dokumentissa, mutta tavallinen käyttäjä ei voi avata turvallisuusasetuksia. QlikView-dokumentit on myös mahdollista tehdä näkymättömäksi offline-käytössä. (Hamsen & García 2012, 529–533; QlikTech International AB 2023.)

QlikView Serverin ja QlikView Windows -asiakaskoneen välinen viestintä on salattua, jos AJAX-asiakaskonetta ei käytetä. Kaikki dokumentit ovat sekoitettuja. Näiden lisäksi QVD-tiedostojen sensitiivistä dataa voidaan salata avainpareilla. (QlikTech International AB 2023.)

4.4.3 Lisensointi

QlikView Desktopista on ilmaiseksi saatavilla henkilökohtainen versio. Henkilökohtaisella versiolla laadittuja dokumentteja ei voi avata, jos joku toinen

käyttäjä on avannut ja tallentanut ne QlikView Serverillä. Toisen käyttäjän dokumenttien avaaminen on estetty QlikView Desktopin henkilökohtaisella versiolla. Henkilökohtaisella versiolla laaditut dokumentit voidaan kuitenkin julkaista QlikView Serverille ja lisensoitu käyttäjä voi avata ne. (Hamsen & García 2012, 28.)

QlikView Server lisensoidaan sarjanumerolla ja valvontanumerolla tai lisensointivaimella. Vuoden 2019 huhtikuusta lähtien QlikView'n asiakkaat voivat käyttää yhtenäistettyä lisenssiä QlikView Server -asennuksille. QlikView Publisher -lisenssi lisää toiminnallisuuksia QlikView Serverille datan päivittämiseen ja jakeluun. (QlikTech International AB 2023.)

Käyttäjän pääsy QlikView Server'illä oleviin dokumentteihin perustuu Client Access License -lisensointiin eli CALeihin, joita on neljää erilaista tyyppiä. CAL voidaan osoittaa yhdelle käyttäjälle (engl. named user CAL), jolloin se on käyttäjän käytössä aina tarvittaessa. Dokumentti CAL (engl. document CAL) sallii yhden nimetyn käyttäjän pääsyn yhdelle dokumentille tarvittavaksi ajaksi. Istuntokohtaisella CALilla (engl. session CAL) voidaan osoittaa nimetylle henkilölle tai tunnistamattomalle käyttäjälle pääsy mille tahansa dokumentille. Istuntokohtaiset CALit varataan kirjautuvien käyttäjien mukaisessa järjestyksessä. Kapasiteettipohjainen CAL (engl. usage CAL) sallii nimetyn tai tunnistamattoman käyttäjän pääsyn dokumentille 60 minuutiksi 28 päivän välisenä aikana. Aikaan pohjautuva CAL osoitetaan kirjautuville käyttäjille sisäänkirjautumisjärjestyksessä. (QlikTech International AB 2023.)

5 TUOTEVERTAILUN VAATIMUKSET

Ennen BI-työkalun valintaa määritetään, millaisia analyysejä tehdään. Sen lisäksi kartoitetaan, millaista on työkalun käyttäjien osaaminen. Vertailuun tehdään selvitys työkalun toimittajasta ja siitä, miten hyvin heidän tarjoamansa tuote sopii organisaation tietoarkkitehtuuriin. Yleensä on turvallista valita suurin toimittaja, vaikka se ei vastaisikaan parhaimmalla tavalla organisaation tarpeita. (Sherman 2014, 166.)

Tuotearviointiin tarvitaan vaatimuksia, joita vasten tuotteita arvioidaan. Vaatimukset määritetään organisaation liiketoiminnallisten ja analyttisten tarpeiden mukaan. Vaatimuksia kerätessä huomioidaan myös datan laatu ja minkälaisista tietolähteistä on kyse. Koko arviointiprosessiin tulisi osallistaa käyttäjiä liiketoiminnan eri aloilta sekä erilaisista käyttäjärooleista. Arvioijat tulisi sitouttaa koko arviointiprosessiin. Vaatimukset priorisoidaan ennen vertailua. Priorisointi tehdään sen mukaan, ovatko vaatimuksien mukaiset ominaisuudet pakollisia, mukava saada, mutta ei pakollisia vai sellaisia, joita ei lopulta oteta käyttöön. Vertailuun otetaan muutama tuote, maksimissaan 3–5 kappaletta. Näiden tuotteiden tulisi olla sellaisia, jotka on mahdollista valita muiden kriteerien mukaan, joita ovat esimerkiksi toimittajan soveltuvuus tai kustannukset. Mukaan voidaan ottaa perinnejärjestelmä (engl. legacy system) vertailun perustason asettamista varten. (Sherman 2014, 55–58, 61, 167–169.)

Liiketoiminnan vaatimukset ja toiminnalliset vaatimukset kerätään haastattelemalla. Haastatteluihin pyydetään osallistujia erilaisista rooleista. Haastattelijan tulee tutustua organisaatioon, ja siihen minkälaisessa roolissa haastateltavat työskentelevät. Haastattelijan tulisi tutustua haastateltavan käyttämiin järjestelmiin ja raportointityökaluihin. Haastateltavilta kysytään, minkälaisesta tiedosta he hyötyisivät, ja miten arvokasta se olisi heille. Haastatteluiden lisäksi voidaan käyttää muita tekniikoita kuten tarinankerrontaa (engl. storyboarding) tai malleja kuten rautalankamalleja ja prototyyppejä. Haastatteluiden pohjalta tunnistetaan tarvittavat tietolähteet, liiketoiminnan prosessit ja sidosryhmät. Haastatteluista tunnistetaan myös toistuvat vaatimukset, ongelmat ja huolet. (Sherman 2014, 56–58, 61.)

On ensiarvoisen tärkeää dokumentoida esille nousevat vaatimukset. Dokumentaatiosta selviää mitkä vaatimukset ovat oleellisia. Osa vaatimuksista voi olla päällekkäisiä tai jopa keskenään ristiriitaisia. Dokumentaation avulla voidaan tunnistaa puutteita vaatimuksissa. Vaatimukset luokitellaan liiketoiminnallisen arvon, liiketoimintaryhmien ja liiketoimintaprosessien mukaan. Vaatimukset luokitellaan myös analyttisten toimintojen mukaan. (Sherman 2014, 55–56, 60.)

Alustavat arviot tehdään tarkastelemalla tuotteiden näyteversioita tai toimittajan näyte-esityksiä. Tuotteille tehdään soveltuvuus selvitys (engl. proof of concept), ja pilotti organisaation omaa dataa hyödyntäen, jos sellainen on mahdollista toteuttaa resurssien puitteissa. Onnistumiskriteerit asetetaan arviointia varten ja määritetään, millaista arvoa vaatimukset tuottavat toteutuessaan. Jokaiselle arvosteltavalle kriteerille annetaan arvosanaskaala, jonka mukaan sitä arvioidaan. Arvostelussa tulee pohtia tuotteen toiminnallisuuksia, sen käytettävyyttä eri käyttäjäroolien kannalta ja teknisestä näkökulmasta. Lisäksi arvioidaan tuotteen soveltuvuutta liiketoiminnan vaatimuksiin. Lopuksi arvosanat lasketaan ja niistä muodostetaan objektiivinen arvio. (Sherman 2014, 167–169.)

Kustannukset

BI-projektin kulut muodostuvat useista muuttuvista ja tuntemattomista tekijöistä. Kustannuksissa tulee arvioida, esimerkiksi ostetaanko tuotteet yksittäisinä tuotteina vai tuotepaketteina. Käyttäjälisensseissä huomioidaan tavalliset käyttäjät ja kehittäjät erikseen ja lisenssien lukumääristä aiheutuvat kustannukset. Muita vaikuttavia tekijöitä ovat esimerkiksi missä ympäristössä tuotetta käytetään. Tuote voi sijaita pilvipalvelussa, tai se voidaan asentaa paikallisesti organisaation omissa tiloissa ja palvelimilla ns. "On Premise". Tuotteeseen liittyvien kulujen lisäksi BI-projektiin liittyy erilaisia henkilöstöresursseihin ja käyttäjien koulutuksiin liittyviä kuluja. Organisaatio voi tarvita oman kehitystiimin tai ostaa ulkoisia konsulttipalveluita. Muita kuluja voivat olla sellaiset, jotka muodostuvat BI-työkalujen ja -alustojen kuten tietovaraston käyttöönotosta, kehittämisestä ja ylläpitämisestä. BI-infrastruktuuriin liittyvät myös tietokantaan tai muihin työkaluihin liittyvät kulut. Kustannuksissa täytyy arvioida, kuinka nopeasti syntyvät kulut katetaan hyödyillä saavutetuilla tuotoilla tai säästöillä, jotta käyttöönotto olisi kannattavaa. Jos organisaatio käyttää perinnejärjestelmää, tulee myös huomioida migraatiokulut perinnejärjestelmästä uuteen järjestelmään. Näihin kuuluvat muun muassa käyttäjien uudelleen kouluttaminen. (Loshin 2012, 21; Sherman 2014, 38, 167–169.)

Toiminnallisuudet

Toiminnallisuuksista esille nousevia toiminnallisia vaatimuksia ovat erilaiset käyttötapaukset, kuten työnkulku dataa analysoitaessa, sekä käyttäjän vuorovaikutus työkalun kanssa. Tarinankerrontatekniikan avulla voidaan tunnistaa miten käyttäjät oikeasti käyttävät työkalua. Kaikkia vaatimuksia ei välttämättä tunnisteta haastattelemalla. Toiminnallisia vaatimuksia voi nousta esille myös takaisinmallintamalla organisaatiossa käytössä olevia BI-ratkaisuja. Vanhat raportointi- ja analyysityökalut eivät ole välttämättä laadukkaasti dokumentoituja, tai niiden dokumentaatio on vanhentunutta. Takaisinmallinnus voidaan tehdä, kun vertaillaan perinnejärjestelmää ja uutta järjestelmää. (Sherman 2014, 52–54, 366.)

Käyttäjäkokemus

Käyttäjäkokemus (engl. UX, user experience) tunnetaan myös käyttäjäkoke-
mussuunnitteluna (engl. UXD, user experience design). Käyttäjäkoke-
mussuunnittelu keskittyy kehittämään käyttäjäystävällisiä kokemuksia, erityisesti digitaalisissa ympäristöissä, kuten internetsivuilla tai ohjelmistoissa. Käyttäjäkoke-
mussuunnittelun keskiössä on käyttäjä, joka on vuorovaikutuksessa oh-
jelmiston kanssa. Hyvin suunnitellussa käyttäjäkokemuksessa otetaan huomi-
oon erityisesti seuraavat avainasiat. Tuotteen tulee olla käyttäjälle hyödyllinen eli se tarjoaa niitä sisältöä tai toimintoja, joita käyttäjä tarvitsee. Tuotteen tulee olla helposti käytettävä, eli se on intuitiivinen ja helppo käyttää. Tuotteen käyt-
täminen ei vaadi erityistä keskittymistä toimintoihin ja sen käyttöprosessi on yksinkertainen, jolloin käyttö on myös nopeaa. Hyvä käyttäjäkokemus on käytettävyyden sekä hyödyllisyyden lisäksi myös miellyttävä. Toisin sanoen oh-
jelma on helposti käytettävä, mutta myös sen ulkoasu ja sisältö ovat puoleen-
savetäviä. Tuotteen tulee olla myös saavutettavissa erilaisille käyttäjäryhmille. (Nichols & Chesnut 2014, 7–11; Canziba 2018, 7–8, 18.)

Laadukas käyttökoke-
mussuunnittelu kasvattaa tuottavuutta, parantaa tuotteen myyntiä. Se myös vähentää koulutuksen ja käyttäjätuen kustannuksia. Käyttäjät ovat tällöin tyytyväisempiä tuotteeseen. (Canziba 2018, 10.)

6 TUOTEVERTAILUN TOTEUTUS

Toimeksiannon mukaisesti tuotevertailu tehtiin Power BI -työkalun ja QlikView-työkalun välillä. Kappaleessa 5 kerrottuja periaatteita sovellettiin tuotevertailun toteutuksessa. Toimeksiantajan vaatimusten painoarvo arvioitiin yhdessä toimeksiantajan kanssa. Arviointiin osallistui henkilöitä, jotka toimivat päällikkö-, esihenkilö- ja asiantuntijatehtävissä.

Kustannukset

Tarkat kustannustiedot on jätetty opinnäytetyön tausta-aineistoon toimeksiantajan käyttöön. Kustannusten arvioitiin muodostuvan seuraavista tekijöistä:

- Lisenssit
- Palvelimet
- Raporttien kehittämistyö ja ylläpito
- Raporttien uudelleen toteutus
- Käyttäjien kouluttautuminen

Lisenssikustannusten katsottiin sisältävän kehittäjälisenssit ja lisenssit raporttien kuluttajille. Raporttien uudelleen toteutus tarkoittaa olemassa olevien raporttien toteutusta toisella työkalulla. Raporttien kehittämistyö ja ylläpito tarkoittaa uusien raporttien laatimista ja ylläpitoa.

Toiminnalliset vaatimukset

Vertailuun valitut vaatimukset on esitetty taulukossa 1. Taulukkoon on merkitty vaatimuksen lisäksi vaatimuksen toteutumisen tärkeys eli painoarvo, jossa 3 = pakollinen, 2 = hyvä, mutta ei pakollinen ja 1 = ei tulisi käyttöön.

Taulukko 1. Tuotevertailussa käytettävät vaatimukset

Kategoria	Vaatus	Painoarvo
Kehitettävyy	Yhteensopivat tietolähteet	3
Kehitettävyy	Datan muokkaus tiedonhaussa	2
Kehitettävyy	Datan tuonti ja aikataulut	2
Kehitettävyy	Kehitystyökalun käytettävyy	2
Kehitettävyy	Tietomallin tarkastelu	1
Kehitettävyy	Tietomallin muokkaus	3
Kehitettävyy	Kehittämisen käyttäjäkokemus	1
Kehitettävyy	Komentokielet	2
Kehitettävyy	Versionhallinta	2
Kehitettävyy	Tietoarkkitehtuuri	3
Kustannukset	Lisensointi	2
Kustannukset	Kehittämiskulut	3
Kustannukset	Siirtyminen uuteen järjestelmään	3
Toiminnallisuudet	Tiimityöskentely	1
Toiminnallisuudet	Raporttien laatiminen	2
Toiminnallisuudet	Datan analysointi työkalussa	3
Toiminnallisuudet	Raporttien tarkastelu	2
Toiminnallisuudet	Modernit teknologiat	1
Käyttäjätkokemus	Responsiivisuus	1
Käyttäjätkokemus	Raporttien vertailu	3
Tietoturva	Tiedon suojaus	3

Taulukosta 1 nähdään, että jokaisesta kategoriasta valittiin pakollisia vaatimuksia. Esimerkiksi tietolähteisiin ja datan käsittelyyn liittyviä pakollisia vaatimuksia olivat yhteensopivat tietolähteet, tietomallin muokkaus ja datan analysointi työkalussa.

Käyttäjätkokemus

Käyttäjätkokemuksen arviointia varten takaisinmallinnettiin olemassa oleva QlikView-dokumentti myös Power BI-työkalulla. QlikView-dokumentissa ja Po-

wer BI -raportissa käytettiin toimeksiantajan laboratoriotiedonhallintajärjestelmän dataa. Käyttäjiltä kerättiin kokemuksia molempien käytettävyydestä. Kokemukset on huomioitu käyttäjäkokemuksen arvioinnissa.

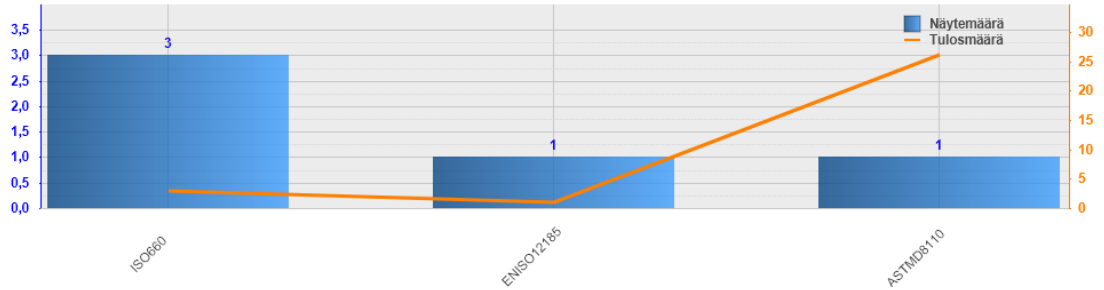
Power BI -raportin toteutus

Power BI -raportin teknistä toteutusta ei selosteta yksityiskohtaisesti opinnäytetyössä tietoturvasyistä. Power BI -raporttia varten tietolähde yhdistettiin Power BI Desktopiin sen tarjoamia liittimiä käyttäen. Ennen tiedonhakua QlikView-dokumentista selvitettiin käytettävät tietokentät, ja varsinainen tiedonhaku tehtiin tietolähteestä SQL-kyselyillä. SQL-kyselyissä haettiin raporttiin ja tietomalliin tarvittavat taulut ja kentät. Jokainen tarvittava taulu haettiin omalla kyselyllä.

Ladattua dataa käsiteltiin Power Query -editorissa. Kaikkien tehtyjen kyselyiden kenttien tietotyypit tarkistettiin ja muutettiin tarvittaessa oikeiksi. Kentät nimettiin uudelleen dataa kuvaavilla nimillä. Power Query -editorissa tehtiin lisää tarvittavia avainkenttiä lisäämällä sarake esimerkin avulla.

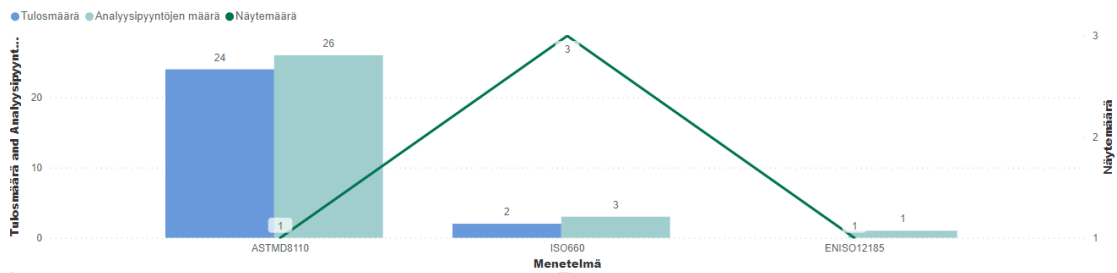
Datan lataamisen jälkeen raportin tietomalli tehtiin luomalla yhteydet taulujen välille Power BI:n tietomallisivulla, raporttinäkymä luotiin QlikView-dokumentin pohjalta ja saadut luvut tarkistettiin QlikView-dokumenttia sekä laboratoriotiedonhallintajärjestelmän dataa vasten. Eroavaisuuksien syyt selvitettiin ja tehdyt muutokset perusteltiin kommenttina raporttiin. Lukujen ja rivien tarkistamisen jälkeen raporttiin tehtiin DAX-kielellä tarvittavia mittareita.

Kuvassa 5 on esitetty kuvankaappaus QlikView-dokumentista, jossa on tietyn ajanjakson näyte- ja tulosmäärä analyysimenetelmien mukaisesti.



Kuva 5. Näyte- ja tulosmäärät analyysimenetelmittäin valittuna ajanjaksona QlikView-dokumentissa

Kuvassa 6 on esitetty kuvankaappaus Power BI -raportista, jossa on tietylle ajanjaksolle valittujen näytteiden, analyysipyyntöjen ja syötettyjen tulosten määrä analyysimenetelmien mukaisesti.



Kuva 6. Näyte-, analyysipyyntö- ja tulosmäärät analyysimenetelmittäin valittuna ajanjaksona Power BI -raportissa

Kuvan 6 Power BI -raportissa kuvan 5 tulosmäärä on eroteltu analyysipyyntöjen ja syötettyjen tulosten määräksi. Raporttien tiedot vastaavat toisiaan. Power BI -raportissa analyysipyyntöjen ja tulosten määrä on tehty palkkikuvaajalla ja näytemäärät on ilmaistu viivakuvaajana.

Työkalujen tarkastelu ja pisteytys

Power BI -työkalun vertailukriteerit ja tulokset on esitetty taulukossa 2. Taulukoon on merkitty mihin vaatimusluokkaan kyseinen rivi kuuluu. Taulukossa on esitetty kunkin kategorian kohdalla arvosana sen mukaan, kuinka hyvin toimiksiantajan tarpeet eli vaatimukset toteutuvat. Lopulta taulukossa on esitetty työkalujen pisteytys. Vaatimusten arvostelu tehtiin arvosana-asteikolla 1–5, jossa 1 = ei toteudu, 3 = toteutuu ja 5 = toteutuu erinomaisesti. Arvosanasta ja painoarvosta on laskettu tulo jokaiselle vaatimukselle. Tarkka lista arvioinnissa

käytetyistä toiminnallisuuksista ja yksityiskohdista on jätetty opinnäytetyön tausta-aineistoksi toimeksiantajan käyttöön.

Taulukko 2. Power BI -työkalun arvostelu

Kategoria	Vaatus	Arvosana	Painoarvo	Tulo
Kehitettävyyys	Yhteensopivat tietolähteet	4	3	12
Kehitettävyyys	Datan muokkaus tiedonhaussa	4	2	8
Kehitettävyyys	Datan tuonti ja aikataulutus	4	2	8
Kehitettävyyys	Kehitystyökalun käytettävyys	5	2	10
Kehitettävyyys	Tietomallin tarkastelu	4	1	4
Kehitettävyyys	Tietomallin muokkaus	4	3	12
Kehitettävyyys	Kehittämisen käyttäjäkokemus	5	1	5
Kehitettävyyys	Komentokielet	5	2	10
Kehitettävyyys	Versionhallinta	3	2	6
Kehitettävyyys	Tietoarkkitehtuuri	4	3	12
Kustannukset	Lisensointi	4	2	8
Kustannukset	Kehittämiskulut	4	3	12
Kustannukset	Siirtyminen uuteen järjestelmään	3	3	9
Toiminnallisuudet	Tiimityöskentely	4	1	4
Toiminnallisuudet	Raporttien laatiminen	4	2	8
Toiminnallisuudet	Datan Analysointi työkalussa	4	3	12
Toiminnallisuudet	Raporttien tarkastelu	4	2	8
Toiminnallisuudet	Modernit teknologiat	4	1	4
Käyttäjäkokemus	Responsiivisuus	4	1	4
Käyttäjäkokemus	Raporttien vertailu	5	3	15
Tietoturva	Tiedon suojaus	5	3	15
Arvosana yhteensä				186

Taulukosta 2 nähdään, että Power BI -työkalun kokonaisarvosana oli 186. Korkein arvosana annettiin pakollisille vaatimuksille raporttivertailun käyttäjäkokemus ja tiedon suojaus. QlikView-työkalun vertailukriteerit ja tulokset on esitetty taulukossa 3. Vaatimukset ja arvostelukriteerit ovat samat kuin taulukossa 2.

Taulukko 3. QlikView-työkalun arvostelu

Kategoria	Vaatus	Arvosana	Painoarvo	Tulo
Kehitettävyyys	Yhteensopivat tietolähteet	3	3	9
Kehitettävyyys	Datan muokkaus tiedonhaussa	2	2	4
Kehitettävyyys	Datan tuonti ja aikataulutus	4	2	8
Kehitettävyyys	Kehitystyökalun käytettävyys	3	2	6
Kehitettävyyys	Tietomallin tarkastelu	4	1	4
Kehitettävyyys	Tietomallin muokkaus	3	3	9
Kehitettävyyys	Kehittämisen käyttäjäkokemus	2	1	2
Kehitettävyyys	Komentokielet	3	2	6
Kehitettävyyys	Versionhallinta	3	2	6
Kehitettävyyys	Tietoarkkitehtuuri	3	3	9
Kustannukset	Lisensointi	2	2	4
Kustannukset	Kehittämiskulut	2	3	6
Kustannukset	Siirtyminen uuteen järjestelmään	3	3	9
Toiminnallisuudet	Tiimityöskentely	2	1	2
Toiminnallisuudet	Raporttien laatiminen	2	2	4
Toiminnallisuudet	Datan Analysointi työkalussa	2	3	6
Toiminnallisuudet	Raporttien tarkastelu	3	2	6
Toiminnallisuudet	Modernit teknologiat	1	1	1
Käyttäjäkokemus	Responsiivisuus	3	1	3
Käyttäjäkokemus	Raporttien vertailu	2	3	6
Tietoturva	Tiedon suojaus	4	3	12
Arvosana yhteensä				122

Taulukosta 3 voidaan todeta, että QlikView-työkalun osalta korkein vaatimuk-
selle annettu arvosana oli 4. QlikView'lle annettu kokonaisarvosana oli 122.
Taulukoiden 2 ja 3 tuloksista voidaan todeta, että Power BI -työkalu on arvos-
telun perusteella soveltuvampi työkalu kuin QlikView.

7 PÄÄTÄNTÖ

Opinnäytetyön tuloksena toteutettiin QlikView- ja Power BI -työkalujen arvi-
ointi. Arviointia varten kerättiin taulukko molempien työkalujen toiminnallisuuk-

sista ja ne pisteytettiin yhdessä toimeksiantajan kanssa. Lopputuloksena voitiin todeta, että Power BI täyttää toiminnalliset vaatimukset paremmin kuin QlikView. Kustannuksien ja kehitettävyyden puolesta Power BI on toimeksiantajalle soveltuvampi. Power BI:ssä saatavilla olevat tekoälyavusteiset toiminnot uupuvat QlikView'stä ja Power BI:n pilvipalvelu on toimeksiantajalle soveltuvampi. Vertailun tuloksena voidaan todeta, että on perusteet korvata QlikView Power BI:lla.

Vertailua varten toteutettiin laboratoriotiedonhallintajärjestelmän datasta Power BI -raportti, jota vertailtiin QlikView-dokumenttiin. Raportin oli tarkoitus esittää samat analyysit kuin QlikView-dokumentissa, ja myös muita toimeksiantajan esittämiä analyysejä. Power BI -raporttiin toteutettiin osa QlikView-dokumentin analyyseistä, sillä tietomallin luominen ja tarkastaminen osoittautui odotettua hankalammaksi. Opinnäytetyön aikana kehitetty Power BI -raportti oli kuitenkin riittävä käyttäjäkokemuksen arviointia varten. Raporttia jatkokehitetään korvaamaan QlikView-dokumentti. Raportin jakelussa hyödynnetään Power BI:n yleisötoimintoa tiedonjakelussa.

Opinnäytetyön tavoite saavutettiin pääosin. Vertailuaineisto on kattava, ja vastaa vertailukysymyksiin. Opinnäytetyössä ei kuitenkaan pysytty aikataulussa. Raporttitoimitus on riittävä, mutta ei saavuta tavoitettaan täysin. Työtä on kuitenkin mahdollista ja suotavaa jatkaa opinnäytetyön jälkeen.

Jos opinnäytetyö tehtäisiin uudestaan, olisi raportin tietomalli suunniteltava huolellisemmin. Vertailusta tulisi kattavampi, jos myös QlikTechin modernimmat työkalut olisivat mukana.

LÄHTEET

- Arnold, J. 2022. Learning Microsoft Power BI. Berkeley: O'Reilly Media, Incorporated. E-kirja. Saatavissa: <https://ebookcentral.proquest.com/lib/xamk-ebooks/detail.action?docID=30101094#> [viitattu 26.9.2023].
- Canziba, E. 2018. Hands-On UX Design for Developers : Design, Prototype, and Implement Compelling User Experiences from Scratch. Birmingham: Packt Publishing, Limited. E-kirja. Saatavissa: <https://ebookcentral.proquest.com/lib/xamk-ebooks/detail.action?docID=5485019#> [viitattu 21.9.2023].
- Hansen, B. & García, M. 2012. QlikView 11 For Developers: Develop Business Intelligence applications with QlikView 11. Birmingham: Packt Publishing, Limited. E-kirja. Saatavissa: <https://ebookcentral.proquest.com/lib/xamk-ebooks/detail.action?docID=1085220> [viitattu 15.2.2024].
- Knight, D., Knight, B., Pearson, M. & Quintana, M. 2018. Microsoft Power BI Quick Start Guide: Build Dashboards and Visualizations to Make Your Data Come Life. Birmingham: Packt Publishing, Limited. E-kirja. Saatavissa: <https://ebookcentral.proquest.com/lib/xamk-ebooks/detail.action?docID=5520883> [viitattu 21.9.2023].
- Loshin, D. 2012. Business Intelligence: the Savvy Manager's Guide. San Diego: Elsevier Science & Technology. E-kirja. Saatavissa: <https://ebookcentral.proquest.com/lib/xamk-ebooks/detail.action?docID=1034439> [viitattu 16.03.2023].
- Maheshwari, A. 2015. Business Intelligence and Data Mining. New York: Business Expert Press. E-kirja. Saatavissa: <https://ebookcentral.proquest.com/lib/xamk-ebooks/detail.action?docID=1911815> [viitattu 8.3.2023].
- Microsoft. 2023a. Create reports and dashboards in Power BI: Design Power BI Reports for accessibility. WWW-dokumentti. Päivitetty 9.3.2023. Saatavissa: <https://learn.microsoft.com/en-us/power-bi/create-reports/desktop-accessibility-creating-reports> [viitattu 29.2.2024].
- Microsoft. 2023b. Data Analysis Expressions (DAX): Aikatieotfunktiot. WWW-dokumentti. Päivitetty 20.10.2023. Saatavissa: <https://learn.microsoft.com/fi-fi/dax/time-intelligence-functions-dax> [viitattu 29.2.2024].
- Microsoft. 2024a. Collaborate, share, and integrate across products with Power BI: Share Power BI reports and dashboards with coworkers and others. WWW-dokumentti. Päivitetty 15.2.2024. Saatavissa: <https://learn.microsoft.com/en-us/power-bi/collaborate-share/service-share-dashboards> [viitattu 29.2.2024].
- Microsoft. 2024b. Get started with Power BI -documentation. WWW-dokumentti. Saatavissa: <https://learn.microsoft.com/en-us/power-bi/fundamentals> [viitattu 29.2.2024].
- Microsoft. 2024c. Power BI Enterprise documentation. WWW-dokumentti. Saatavissa: <https://learn.microsoft.com/en-us/power-bi/enterprise> [viitattu 29.2.2024].

Microsoft. 2024d. Power BI:n hinnoittelu. WWW-dokumentti. Saatavissa: <https://powerbi.microsoft.com/fi-fi/pricing/> [viitattu 29.2.2024].

Microsoft. 2024e. Power BI mobile apps: What are the Power BI mobile apps?. WWW-dokumentti. Päivitetty: 1.1.2024. Saatavissa: <https://learn.microsoft.com/en-us/power-bi/consumer/mobile/mobile-apps-for-mobile-devices> [viitattu 29.2.2024].

Microsoft. 2024f. The Power BI service for business users: Tutorial – Getting around in Power BI Service. WWW-dokumentti. Päivitetty 9.2.2024. Saatavissa: <https://learn.microsoft.com/en-us/power-bi/consumer/end-user-experience> [viitattu 29.2.2024].

Microsoft. 2024g. Transform, shape, and model data in Power BI. WWW-sivu. Saatavissa: <https://learn.microsoft.com/en-us/power-bi/transform-model/> [viitattu 28.2.2024].

Nichols, KP. & Chesnut, D. 2014. UX for Dummies. Newark: John Wiley & Sons, Incorporated. E-kirja. Saatavissa: <https://ebookcentral.proquest.com/lib/xamk-ebooks/detail.action?docID=1674223#> [viitattu 21.9.2023].

QlikTech International AB. 2023. Documentation: QlikView. WWW-dokumentti. Saatavissa: https://help.qlik.com/en-US/qlikview/May2023/Content/QV_HelpSites/Home.htm [viitattu 15.2.2024].

Sherman, R. 2014. Business Intelligence Guidebook: From Data Integration to Analytics. Elsevier Science & Tehnology. E-kirja. Saatavissa: <https://ebookcentral.proquest.com/lib/xamk-ebooks/detail.action?docID=1832704> [viitattu 23.05.2023].