



Tietolähteiden yhdistäminen Power BI:ssa

Jesse Majaranta

OPINNÄYTETYÖ
Marraskuu 2023

Tietojenkäsittelyn tutkinto-ohjelma
Ohjelmistotuotanto

TIIVISTELMÄ

Tampereen ammattikorkeakoulu
Tietojenkäsittelyn tutkinto-ohjelma
Ohjelmistotuotanto

MAJARANTA, JESSE:
Tietolähteiden yhdistäminen Power BI:ssa

Opinnäytetyö 36 sivua
Marraskuu 2023

Tavoitteena opinnäytetyössä on yhdistää Power BI:ssa kaksi eri tietolähdettä ja luoda niihin pohjautuva raportti. Raportin on tarkoitus auttaa havainnoimaan myyjien kiviporalaiteiden osien määrää suhteessa jo asiakkaan omistamiin laitteisiin. Tällä tavalla pystytään havaitsemaan asiakkaita, jotka tekevät vähemmän ostoja kuin muut.

Opinnäytetyön toimeksiantaja on Sandvik Mining and Construction Oy, joka pääasiassa valmistaa kallioporauslaitteita. Työ tehdään Parts & Service -divisioonassa, joka on keskittynyt jälkimarkkinaan myymällä varaosia ja huoltopalveluita.

Teoriaosuudessa tutkitaan Power BI:ta palveluna. Siinä selvitetään, mitä raportin luomiseen vaaditaan ja miten se toteutetaan. Alussa aihetta pohjustetaan business intelligenen historialla ja muilla Power BI:ta vastaavilla palveluilla.

Tietolähteitä ei saatu yhdistettyä tavalla, joka olisi tuottanut toivotun lopputuloksen. Ne saatiin kuitenkin yhdistettyä tavalla, jolla raporttiin saatiin valmiiksi taulukko, joka näyttää käyttäjälle asiakkaan omistamat ja 12 kuukauden aikana ostetut tuotteet ilman alakategorioita. Näiden arvojen avulla pystyttiin myös laskemaan niiden välinen suhde asiakasta kohden ja keskiarvo maassa.

ABSTRACT

Tampereen Ammattikorkeakoulu
Tampere University of Applied Sciences
Degree Programme in Business Information Systems
Software development

MAJARANTA, JESSE:
Combining Data Sources in Power BI

Bachelor's thesis 36 pages
November 2023

The goal of the thesis is to merge two data sources in Power BI and create a report with them. The purpose of the report is to help observe the number of parts of rock drilling equipment sold in relation to the equipment already owned by the customer. In this way, it is possible to detect customers who make fewer purchases than others.

The client of this thesis is Sandvik Mining and Construction Oy who mainly produces rock drills. The thesis was done for the Parts & Services division, which is focused on the aftermarket by selling spare parts and maintenance services.

Power BI is inspected as a service in the theory part of the thesis. The theory part explains what is needed to create a report and how it is done. The topic is also connected to the history of business intelligence and other services similar to Power BI.

The data sources used could not be linked in a way that would produce the desired result. However, they were combined in a way that a table was completed for the report. It shows the user the products owned by the customer and purchased during the past 12 months without subcategories. With these values, it was possible to measure the ratio between those numbers and the average national ratio.

Key words: Power BI, data, visualization

SISÄLLYS

1	JOHDANTO	7
2	BUSINESS INTELLIGENCE	8
	2.1 Mitä tarkoittaa business intelligence?	8
	2.2 BI:n evoluutio	8
	2.2.1 Tiedon määrän kasvu 1960- ja 1970-luvuilla	8
	2.2.2 Raportoinnin haasteet 1990-luvulla	9
	2.2.3 Self-service BI	9
	2.3 Miksi BI on tärkeää	10
	2.4 BI-työkalut	10
	2.4.1 BI-työkalujen toiminta yleisesti	10
	2.4.2 Tableau Software	11
	2.4.3 Mikä on Qlik ja Qlik Sense?	12
3	POWER BI:N TOIMINNAT JA KÄYTTÖ	13
	3.1 Mikä on Power BI?	13
	3.2 Power BI:n käyttö	13
	3.2.1 Näkymävaihtoehdot Power BI:ssa	13
	3.2.2 Kuinka data lisätään?	14
	3.2.3 Datan muuntaminen	16
	3.2.4 Taulukoiden väliset suhteet	18
	3.2.5 Mittarit ja lasketut sarakkeet	18
	3.2.6 Visualisointien luominen	19
4	DATAN VISUALISOINTI	20
	4.1 Datan visualisoinnin hyödyt yleisesti	20
	4.2 Datalla voi kertoa tarinan Power BI:ssa	21
	4.3 Parhaat käytännöt	22
5	ENNEN RAPORTIN TOTEUTUSTA	23
	5.1 Taustatietoa työstä	23
	5.2 Palvelun vaihtuminen uuteen	24
	5.3 Mikä ongelma ratkaistaan?	25
6	RAPORTIN TOTEUTUS	26
	6.1 Datan lisääminen	26
	6.2 Taulukoiden väliset yhteydet	26
	6.3 Tietolähteiden väliset ongelmat	29
	6.3.1 Taulukoiden linkittäminen	29
	6.3.2 Ongelmanratkointia yhteyden löytämiseksi	30
	6.3.3 Lopullinen datamalli	31

6.4 Lopputulos	32
7 POHDINTA	34
LÄHTEET	35

LYHENTEET JA TERMIT

BI	Business intelligence, liiketoimintatiedot
DAX	Data Analysis Expression, funktiokirjasto
DSS	Päätöksenteontukijärjestelmä, ensimmäinen tiedonhallintajärjestelmä
IT	Information technology, tietotekniikka
KPI	Key performance indicator, suorituskykyä mittaava luku
PBIX	Power BI:n työpöytäsovelluksen tiedostomuoto
SaaS	Software as a Service, internetin välityksellä toimiva ohjelmisto
Tietojoukko/Dataset	Power BI:n tietomalli, sisältää tiedot ja taulukoiden välisen rakenteen

1 JOHDANTO

Tässä opinnäytetyössä tutkitaan Power BI -raportin rakentamista tietolähteen valitsemisesta visualisointiin. Selvitetään, mitä eri vaiheissa tapahtuu ja miten ohjelmistoa niissä käytetään. Teoriaosuudessa vastataan kysymykseen, mitä tarkoittaa termi business intelligence. Osioon sisältyy myös historiaa siitä, miten dataa on hyödynnetty liiketoiminnassa vuosikymmenien ajan.

Työssä käydään läpi, mitä Power BI -raportin luomiseen kuuluu. Ennen visualisointien luomista tulee lisätä data, mahdollisesti muovata sitä ja luoda taulukoiden välille suhteet. Tämän jälkeen selvitetään, miten data tulisi näyttää käyttäjälle.

Työn toimeksiantaja on Sandvik Mining and Construction Oy. Se valmistaa pääasiassa kaivosporalaitteita. Käytännön työssä tavoitteena on toteuttaa Power BI:lla raportti, jossa yhdistetään kaksi eri tietolähdettä. Ensimmäinen tietolähde sisältää datan asiakkaiden jo omistamista kaivosporalaitteista ja toinen myydyistä kaivosporista sekä niiden osista. Kun nämä yhdistetään, pystytään selvittämään, kuinka paljon asiakkaat ovat ostaneet osia suhteessa jo omistamiinsa laitteisiin. Tarkoituksena on saada näkyviin asiakkaat, jotka ostavat suhteessa muita asiakkaita vähemmän. Tämän avulla myyntiosasto tietää, kenelle tulisi myydä enemmän tuotteita.

2 BUSINESS INTELLIGENCE

2.1 Mitä tarkoittaa business intelligence?

Business intelligence -käsitteellä tarkoitetaan ohjelmistoa, joka muuttaa yrityksen liiketoimintaan liittyvän datan käyttäjälle ymmärrettävään muotoon. Helposti ymmärrettävien visuaalisten havaintojen pohjalta voidaan tehdä strategisia ja taktisia päätöksiä. Havaintoja voidaan esittää erilaisissa visuaalisissa muodoissa, kuten esimerkiksi kaavioina, taulukoina, raportteina tai koontinäyttöinä. Nämä havainnot ovat nopeasti saatavilla sekä ymmärrettävissä. Business intelligence kääntyy suomeksi sanaksi liiketoimintatiedot, mutta yleisesti se tunnetaan lyhenteellä BI. (Mitä ovat liiketoimintatiedot? n.d.; What is business intelligence n.d.; Pratt 2012.)

BI ei kerro suoraan käyttäjälle, mitä tulisi tehdä tai mitä tulevaisuudessa tulee tapahtumaan. Se tuo ilmi, mitä data kertoo kyseisellä hetkellä sekä sen, mitä on tapahtunut aikaisemmin. Tarkoitus on auttaa käyttäjää löytämään merkityksellinen data, jonka pohjalta voidaan tehdä havaintoja ja liiketoimintaa edistäviä päätöksiä. (Pratt 2012.)

2.2 BI:n evoluutio

2.2.1 Tiedon määrän kasvu 1960- ja 1970-luvuilla

Ensimmäisen kerran termiä business intelligence käytti Richard Millar Devens vuonna 1865. Hän siteerasi pankkiiria nimeltä Sir Henry Furnese, joka keräsi tietoa ja käytti sitä hyödykseen kilpailijoita vastaan. (Foote 2023.)

Tiedon määrä kasvoi paljon 1960- ja 1970-luvuilla. Sen takia tarvittiin jotain, millä datan sai pidettyä järjestyksessä ja hyvin tallessa. Silloin kehitettiin ensimmäiset tiedonhallintajärjestelmät. (What is business intelligence n.d.)

Edgar Codd ehdotti relaatiotietokannan kehittämistä vuonna 1970. Ensimmäinen tiedonhallintajärjestelmä, joka kehitettiin, oli päätöksenteontukijärjestelmä eli DSS. DSS on tietokoneohjelma, jota käytetään yrityksissä apuna päätöksenteossa. Ohjelmaan syötetään dataa ja vastaavasti se antaa vastauksen parhaista vaihtoehdoista. Nykyaikaisen BI:n sanotaan kehittyneen DSS-tietokannasta. (What is business intelligence n.d.; Definition decision support system (DSS) 2021; Foote 2023.)

2.2.2 Raportoinnin haasteet 1990-luvulla

1990-luvun alussa BI:n suosio kasvoi paljon. Yritykset loivat staattisia raportteja, joilla mitattiin KPI-lukemia ja suorituskykyä. Teknologian ollessa vielä monimutkaista tietomäärän kasvu loi yrityksille haasteita. Se johti esimerkiksi siihen, että organisaatioiden sisällä raportoinnissa saattoi tulla ilmi ristiriitaisia tuloksia. Tämän takia organisaatioihin perustettiin erillinen BI-työryhmä yleensä IT-osaston sisälle. (What is business intelligence n.d.; Tableau 2023.)

Raportteja työstettiin samalla tavalla, kuin normaaleja IT-projekteja. Raportin pyynnöstä IT-osastolle sen valmistumiseen kesti yleensä liian kauan. Raportit myös valmistuivat usein myöhässä. Vaikka BI-osaamisen keskittäminen IT-osastolle toi toivotunlaisia tuloksia, toiminta ei ollut tarpeeksi ketterää ja BI:n vaikutukset jäivät vähäisiksi. Muuta vaihtoehtoa tähän aikaan ei ollut, koska datan analysointiin vaadittiin niin paljon osaamista, että kuka tahansa ei sitä voinut suorittaa. (What is business intelligence n.d.; Tableau 2023.)

2.2.3 Self-service BI

Nykyään suurin osa BI-työkaluista on saatavilla kenelle tahansa. Enää ei tarvitse olla IT-alan ammattilainen, jotta pystyy käyttämään dataa hyödyksi päätöksenteossa. Tätä mallia kutsutaan self-service BI:ksi, eli se on BI:n itsepalvelumalli. Yrityksissä tarvitaan enemmän ketterää toimintaa, ja tämän takia self-service BI on suosituin raportoinnin muoto tällä hetkellä. (Pratt 2019; Tableau 2023.)

BI-työkalujen käyttö ei tietenkään ole niin helppoa, että se onnistuisi ilman harjoittelua. Työkalujen käyttöliittymät on rakennettu kuitenkin niin helppokäyttöiseksi, että kaikki työntekijät oppivat niiden käytön halutessaan. BI-työkaluja käytettäessä tulee käyttäjälle myös enemmän ymmärrystä niiden tarjoamista toiminnoista ja hyödyistä. Se rohkaisee uusia käyttäjiä luomaan raportteja entistä enemmän. (Pratt 2019.)

Self-service-mallilla on tietenkin myös negatiivisia vaikutuksia raportointiin. Jos organisaatiossa moni henkilö tekee BI-raportteja, ne todennäköisesti ovat tyyliittäin keskenään erilaisia. Niissä ei myöskään ole otettu huomioon seikkoja, joista ammattilaiset ovat tietoisia. Yhtenäisillä linjauksilla ja harjoituksilla yrityksen raportoinnista saadaan todella hyvä. (Pratt 2019.)

2.3 Miksi BI on tärkeää

BI:n avulla yritykset saavat tärkeää tietoa niiden tilanteesta. Datan analysointi tapahtuu nopeasti, mikä tekee ongelmiin reagoimisen helpommaksi. Oli tilanne mikä tahansa, datan avulla voidaan tehdä päätöksiä siitä, mitä tehdään tulevaisuudessa. BI auttaa yrityksiä saavuttamaan kilpailuedun, kun se tehdään oikein. (Mitä ovat liiketoimintatiedot? n.d.; Tableau 2023.)

Myyntityössä voidaan käyttää BI:tä havaitsemaan asiakkaan käytöstä tai ostotrendejä. Saadaan siis tietoon, mitä tuotteita ihmiset haluavat ostaa. Talouspuolella voidaan tarkastella nykyistä sekä mennyttä tilannetta. Tämä auttaa enustamaan tulevaisuuden trendejä. (Mitä ovat liiketoimintatiedot? n.d.)

2.4 BI-työkalut

2.4.1 BI-työkalujen toiminta yleisesti

BI-työkalut, joilla raportteja luodaan, toimivat usein samalla periaatteella. Alustat yleensä saavat datan tietovarastoista. Tietovarastot säilövät dataa monesta eri

lähteestä. Ne kokoavat tiedot nouto, muunto ja lataus -menetelmän avulla. Lopputuloksena on uudelleenmallinnettu ja yhtenäinen tietojoukko. Tätä tietojoukkoa käytetään datana raportointiin ja siitä voidaan tehdä erilaisia visualisointeja, kuten esimerkiksi taulukoita ja kaavioita. (Mitä ovat liiketoimintatiedot? n.d.; What is business intelligence? n.d.)

BI-työkaluja on paljon erilaisia. CIO listasi niistä 12 suosituinta, joidenka joukossa on Power BI (Sayer & Olavsrud 2021). Sen lisäksi huomasin kaksi muuta tutun kuuloista analytiikkatyökalua, Tableau ja Qlik, joista otin selvää enemmän.

2.4.2 Tableau Software

Tableau Software perustettiin vuonna 2003. Sen tavoitteena on tuoda data ja sen analysointi saavutettavaksi kaikille visualisoinnin kautta. Yritys haluaa auttaa asiakkaita näkemään ja ymmärtämään dataa. (What is Tableau? 2023.)

Tableaun Softwaren oma teknologia VizQL patentoitiin, kun se kehitettiin. Se esittää käyttäjälle datan helppossa käyttöliittymässä, joka hyödyntää raahaa ja pudota -toimintoa. Yritys haluaa kehittää työkalujaan, jotta käyttäjät saisivat datasta vastauksen helposti ja nopeasti. Tavoitteena on myös tuoda esiin havaintoja, joita käyttäjä ei välttämättä itse huomaisi. Kehityksen kohteena ovat myös koneoppimisen, tilastitiikan, selkokielen ja älykkään datan hyödyntäminen (What is Tableau? 2023.)

Tableau on monikäyttöinen työkalu. Sitä voi käyttää paikallisesti tietokoneella tai Tableaun palvelimella. Kohderyhmänä ovat suuret ja keskisuuret yritykset. Sitä voidaan käyttää melkein pä minkälaisessa yrityksessä tahansa terveydenhuollosta taloushallintoon. Tableau Software tarjoaa kustomoituja versioita palvelustaan eri alojen yrityksille. Tableau mahdollistaa myös kyselyn luomisen pelkkää selkokieltä käyttämällä. Käyttäjä syöttää tekstisyötteenä, mitä hän haluaa nähdä, ja tekoäly pyrkii vastaamaan tähän kysymykseen tekemällä kyselyn ja esittämällä sen käyttäjälle. (Sayer & Olavsrud 2021.)

2.4.3 Mikä on Qlik ja Qlik Sense?

Qlikin tavoitteena on, että jokainen yrityksessä pääsee dataan käsiksi tiettyyn pisteeseen asti. Kaikki data soveltuu suurimpiin osiin tietokannoista. Qlikin mukaan kuitenkin yrityksen ohjelmisto toimii monen erilaisen tiedon kanssa ja se auttaa löytämään yhteyden myös erilaisten datamuotojen välille. (Qlik Sense® | Modern Cloud Analytics n.d.; Sayer & Olavsrud 2021.)

Qlik Sense -analytiikkatyökalulla pystyy tekemään valintoja vapaasti ja nopeasti. Associative Engine käyttää hyödyksi tekoälyä. Ne esimerkiksi tarjoavat käyttäjälle suosituksia havainnoista, selkokielellä hakemisen ja ennustavan analyysin. Qlik Sense on saatavana paikallisena sekä pilviversiona (Qlik Sense® | Modern Cloud Analytics n.d.; Sayer & Olavsrud 2021.)

3 POWER BI:N TOIMINNAT JA KÄYTTÖ

3.1 Mikä on Power BI?

Power BI on Microsoftin kehittämä ympäristö liiketoiminnan analysointiin. Sillä pystyy tekemään visuaalisia raportteja paikallisista sekä pilvessä sijaitsevista tietolähteistä. Power BI -raportit pystytään yhdistämään helposti muihin Microsoftin sovelluksiin. (Mikä on Power BI? n.d.)

Power BI:lla pystyy tekemään visuaalisia raportteja monesta eri tietolähteestä. Usean eri tietolähteen pystyy yhdistämään yhdeksi kokonaisuudeksi. Dataa pystyy muovaamaan ja siistimään haluamakseen. (Mikä on PowerBI? n.d.)

Ilmaisella versiolla pystyy tekemään raportteja paikallisesti. Pro-version avulla raportit on mahdollista jakaa muille käyttäjille ja yhdistää muihin Microsoftin 365-palveluihin. Premium-versio, joka on tarkoitettu suurille yrityksille, auttaa siivoamaan ja muokkaamaan dataa, joka on säilötyinä esimerkiksi Microsoftin Dynamics 365:ssä, Azuressa tai kolmannen osapuolen palvelussa. (Sayer & Olavsrud 2021.)

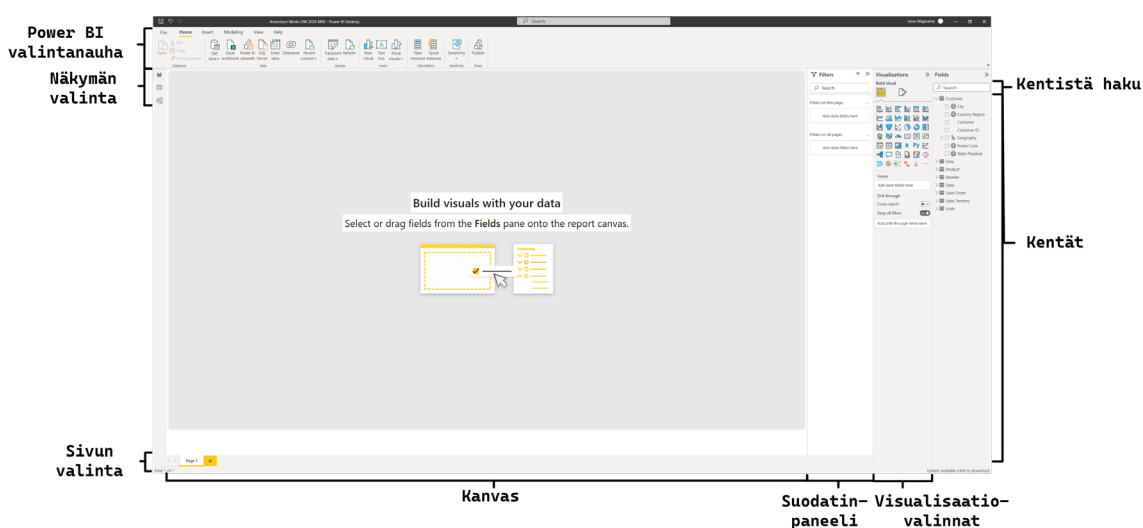
3.2 Power BI:n käyttö

3.2.1 Näkövaihtoehdot Power BI:ssa

Kun Power BI:n työpöytäsovelluksen avaa, tulee aina ensimmäiseksi näkyviin raporttinäkymä (kuva 1). Sitä käytetään aina raporttien luomiseen. Tässä näkyvässä raporttiin voidaan lisätä visualisointeja, jotka käyttävät raporttiin lisättyä dataa. Ennen tätä vaihetta on käyttäjän pitänyt ladata data ja mahdollisesti siivota sekä muotoilla sitä haluamakseen. (Aspin 2022.)

Kuvaan 1 on merkitty raporttinäkymän työkalut ja valinnat. Ylhäällä valintanauhasta löytyy tärkeät työkalut raportin tekemiseen. Näkymän valinnasta pystyy

vaihtamaan kolmen eri näkymätyyppin välillä. Ensimmäisenä on raporttinäkymä, joka on kuvassa valittuna. Kaksi muuta näkymää ovat datanäkymä ja mallinnusnäky-
 näkymä. Datanäkymässä voi muokata dataa ja tehdä laskelmia. Mallinnusnäky-
 mässä voi yhdistää eri tietolähteitä sekä taulukoita toisiinsa. Keskellä raport-
 tinäkymässä (kuva 1) on kanvas, johon luodaan visualisoinnit, kuten taulukot ja
 kaaviot. Alhaalta vasemmalta löytyy sivunvaihtopainikkeet. Oikealla kuvasta löy-
 tyy suodatinpaneeli, visualisointien valinnat ja kentät. Suodatinpaneelistä pystyy
 suodattamaan valittua dataa.

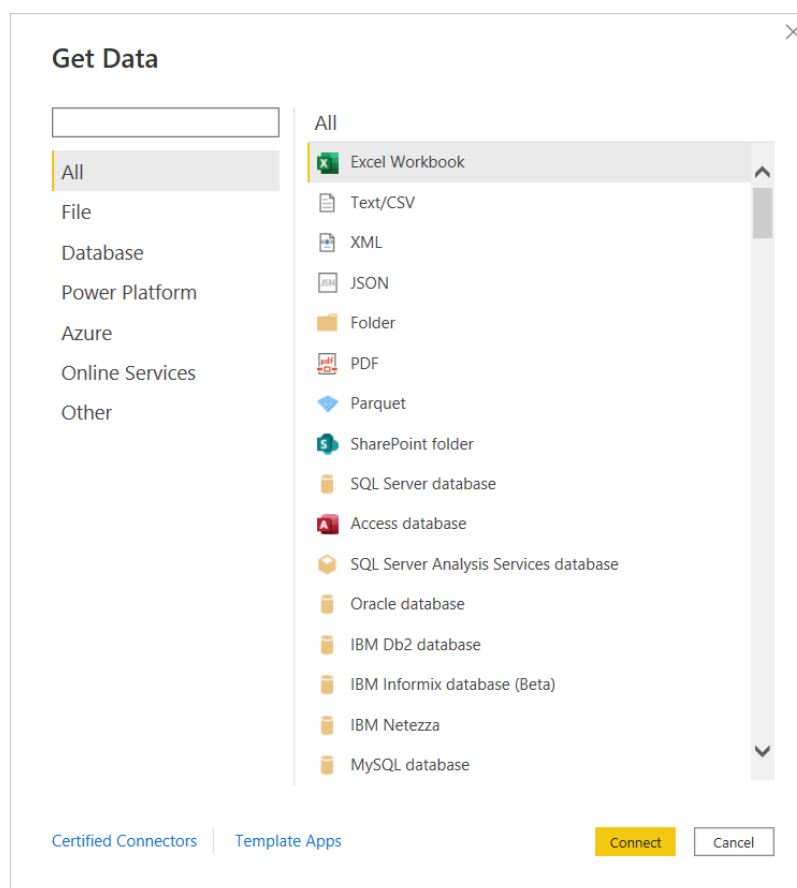


KUVA 1. Power BI: työpöytäsovelluksen oletusnäky.

3.2.2 Kuinka data lisätään?

Power BI:hin pystyy lisäämään dataa monesta erilaisesta lähteestä. Useamman tietolähteen pystyy yhdistämään yhdeksi kokonaisuudeksi. (Mikä on Power BI? n.d.) Lähteenä voi toimia paikallinen tiedosto, SaaS-sovellus tai tietokanta (Larson 2020).

Tietolähteen lisääminen tapahtuu Power BI:n työpöytäsovelluksessa. Se tehdään painamalla painiketta, jossa lukee Get data. Ruudulle aukeaa valikko, josta voi valita datan lähteen (kuva 2). Valikossa on monta vaihtoehtoa, ja käyttäjän tulee tietää, mitä niistä tulee käyttää. (Larson 2020.)



KUVA 2. Tietolähdevalikko Power BI:ssa.

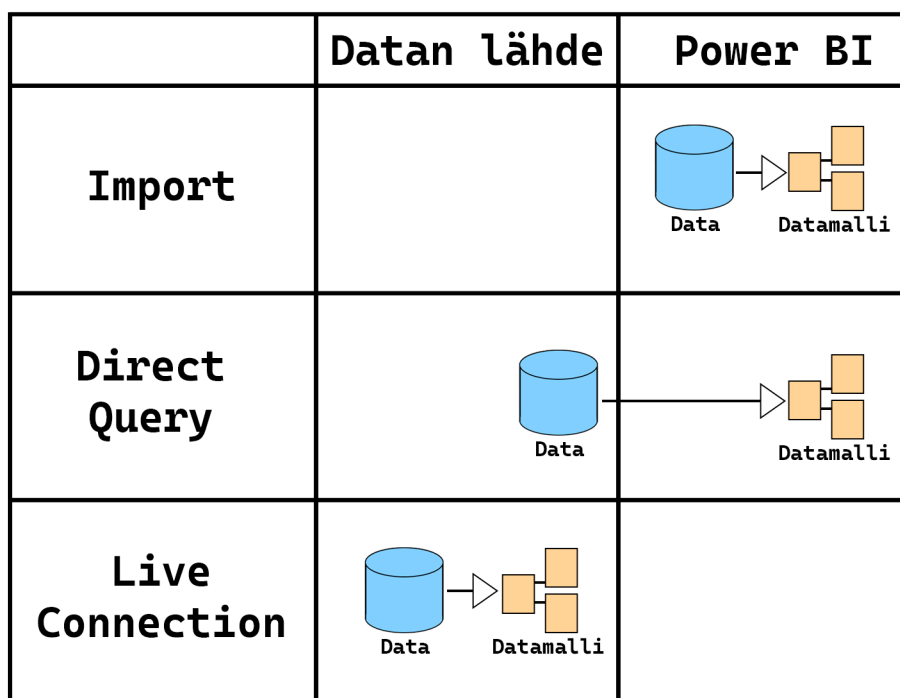
Tietolähde on mahdollista yhdistää Power BI:hin kolmella eri tavalla. Ne ovat englanniksi Import, Live Connection ja Direct Query. (Larson 2020.) Kuvassa 3 havainnollistetaan datan ja datamallin sijainnit eri yhdistämistapoja käytettäessä.

Käyttäessä vaihtoehtoa Import tietolähteestä haettu data tallennetaan Power BI:n PBIX-tiedostoon. Tällöin Power BI käyttää dataa, joka on paikallisesti tallennettuna tietokoneella. Jos datan haluaa päivittää, sovellus ottaa yhteyttä datan lähteeseen ja päivittää muuttuneet tiedot PBIX-tiedostoon. Koska tiedot on tallennettu paikalliseen muistiin, tietojen hakeminen on tehokasta. Taulukoita ja niiden välisiä suhteita pystyy mallintamaan helposti. Paikallisesti tallennetussa datassa huonona puolena on se, että tieto voi olla vanhentunutta, jos sitä ei ole päivittänyt. (Larson 2020.)

Datan voi hakea suoraan tietolähteestä Live Connection -vaihtoehtoa käyttäen. Tätä voi käyttää, kun lähteenä on esimerkiksi SQL-tietokanta, Azure Analysis Services tai Power BI -tietojoukko. Tällöin Power BI tekee kyselyn tiedon lähteeseen ja hakee datan, kun sitä tarvitaan. Dataa ei tarvitse mallintaa itse, vaan se

on jo valmiina. Tämä vaihtoehto on nopea tapa hyödyntää valmiiksi mallinnettua dataa. Koska data on jo valmiiksi mallinnettu, tietoa ei kuitenkaan voi muokata paljoakaan, mikä on tämän tavan heikkous. (Larson 2020.)

Direct Query -vaihtoehdon ominaisuudet ovat Importin ja Live Connectionin välillä. Data tulee suoraan lähteestä ilman, että se tallentuu Power BI -malliin. Tieto pitää paikkaansa, koska se päivittyy sitä mukaa, kun sitä tarvitaan. Se ei kuitenkaan ole aina tehokasta, sillä jatkuvat kyselyt ja suuri määrä dataa vaativat paljon suorituskykyä. Laskelmien tekeminen voi olla tästä syystä hidasta. Datan malli kuitenkin on Power BI:ssa paikallisesti, joten sitä voi mallintaa haluamallaan tavalla. (Larson 2020.)



KUVA 3. Datan ja datamallin sijainnit eri vaihtoehdoissa.

3.2.3 Datan muuntaminen

Kun datan hakee Import-ominaisuudella, sen sisältöä ja rakennetta pystyy muokkaamaan. Tällöin sen saa vastaamaan parhaiten omaa käyttötarkoitusta. Jos dataa on jotain vialla, sen huomaa, ennen kuin alkaa tekemään raporttia. Tämä tehdään Power BI:n Power Query -editorilla. (Larson 2020.)

Power Query -editori on hyödyllinen erilaisissa tilanteissa. Sillä voi esimerkiksi poistaa taulukosta sarakkeita, joita ei tarvita. Kaikki turha data kannattaa poistaa, koska se vain aiheuttaa sotkua. Datan muotoilua voi myös muuttaa haluamukseen. Esimerkiksi on mahdollista vaihtaa desimaalien määrää tai päivän muotoilua haluamukseen. (Larson 2020.)

Kuvassa 4 näkyy Power Queryn oletusnäkömä. Se on jaettu neljään osaan. Vasemmalla näkyvät kyselyt ja oikealla näkyvät niihin tehdyt muutokset. Keskellä näkyy data valitun muutoksen jälkeen. Oletuksena valittuna oikealla olevasta pallosta on viimeisin muutos eli listan alin kohta. Tällöin dataan otetaan huomioon kaikki siihen tehdyt muutokset. Valitsemalla listasta tehdyn muutoksen pystyy katsomaan, miten se on vaikuttanut dataan. Sen asetuksia pystyy myös muokkaamaan jälkikäteen. Listattujen muutosten järjestystä pystyy vaihtamaan raa haamalla hiirellä valittua vaihtoehtoa. Ylhäällä ruudun poikki kulkevasta valintanauhasta pystyy valitsemaan muutoksen, jolla haluaa muokata dataa. (Larson 2020.)

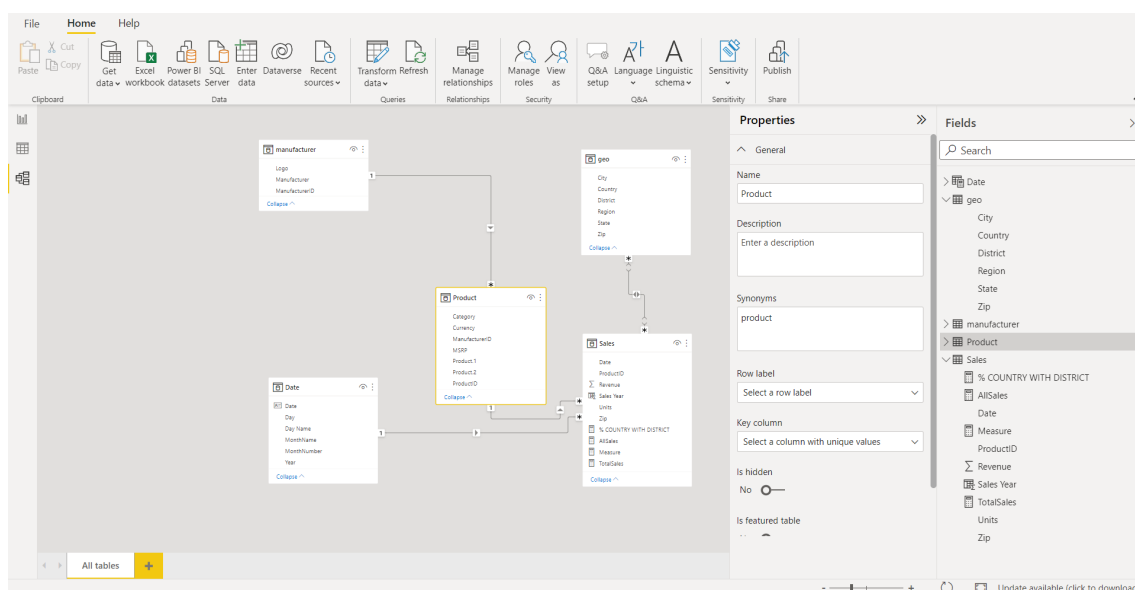
ProductID	Product	Product.2	Category	ManufacturerID	Currency	MSRP
1	Abbas MA-01	All Season			USD	412,13
2	Abbas MA-02	All Season	Mixture		USD	329,78
3	Abbas MA-03	All Season			USD	963,38
4	Abbas MA-04	All Season			USD	828,98
5	Abbas MA-05	All Season			USD	745,5
6	Abbas MA-07	All Season			USD	451,45
7	Abbas MA-06	All Season			USD	329,78
8	Abbas MA-08	All Season			USD	485,89
9	Abbas MA-09	All Season			USD	634,73
10	Abbas MA-10	All Season			USD	681,98
11	Abbas MA-11	All Season			USD	761,25
12	Abbas MA-12	All Season			USD	456,7
13	Abbas MA-13	All Season			USD	456,7
14	Abbas MA-14	All Season			USD	419,95
15	Abbas MA-15	All Season			USD	472,45
16	Abbas MA-16	All Season			USD	711,38
17	Abbas MA-17	All Season			USD	414,75
18	Abbas MA-18	All Season			USD	393,7
19	Abbas MA-19	All Season			USD	393,7
20	Abbas MA-20	All Season			USD	509,25
21	Abbas MA-21	All Season			USD	490,5
22	Abbas MA-22	All Season			USD	551,25
23	Abbas MA-23	All Season			USD	758,63
24	Abbas MA-25	All Season			USD	792,75
25	Abbas MA-26	All Season			USD	847,88
26	Abbas MA-24	All Season			USD	329,78
27	Abbas MA-27	All Season			USD	446,2
28	Abbas MA-28	All Season			USD	431.6125

KUVA 4. Power Query -editorin oletusnäkömä.

3.2.4 Taulukoiden väliset suhteet

Usein Power BI:ta käyttäessä ladattuna on useamman taulukon tiedot. Power BI yleensä tunnistaa taulukoiden väliset suhteet automaattisesti sarakkeiden nimien perusteella. Taulukoita, joiden tiedoista tehdään kysely, tarvitsee olla tällöin kaksi tai useampi. Käyttäjän ei tällöin tarvitse tehdä mitään. Kuitenkin välillä suhteita joutuu luomaan itse. (Suhteiden luominen ja hallinta Power BI Desktopissa 2023.)

Malli-näkymässä (kuva 5) pystyy muokkaamaan taulukoiden välisiä suhteita. Suhde näkyy viivana taulukoiden välillä. Jokaisella suhteella on kardinaliteetti. Sen arvo voi olla monta yhteen (*:1), yksi yhteen (1:1), yksi moneen (1:*) tai monta moneen (*:*). Asetus kertoo, montako kertaa arvo voi esiintyä kummassakin taulukossa. Yleisin asetus on monta yhteen (*:1), jossa arvo voi esiintyä toisessa taulukossa useammin kuin kerran ja hakutaulukossa vain kerran. (Suhteiden luominen ja hallinta Power BI Desktopissa 2023.)



KUVA 5. Esimerkkikuva Malli-näkymästä.

3.2.5 Mittarit ja lasketut sarakkeet

Tärkeä osa ennen datan visualisointia on Power BI:n mittarit ja lasketut sarakkeet. Niiden avulla saa tuotua datasta kaikki havainnot esiin. Mittarit lasketaan

Power BI:n DAX-kielellä. DAX toimii lähes samalla tavalla kuten Excel funktiot, mutta se on kuitenkin tarkoitettu suhteellisten tietojen käsittelyyn. DAX-kirjastossa on yli 200 funktiota, operaattoria ja rakennetta. (Larson 2020; Suhteiden luominen ja hallinta Power BI Desktopissa 2023.)

Yksinkertaisimmillaan mittarit ovat laskutoimituksia, kuten summat, keskiarvot ja määrät. Ne muuttuvat aina raportin mukana tietojen päivittyessä, jotta ne ovat ajan tasalla. Niitä pystyy luomaan raportti-, data- ja mallinäkyessä. (Suhteiden luominen ja hallinta Power BI Desktopissa 2023.)

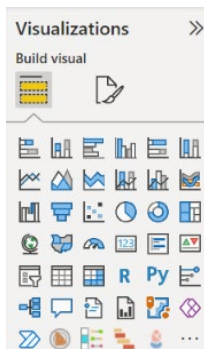
Kuvassa 6 on esimerkki DAX-kielellä muodostetusta mittarista. Sen avulla on laskettu myynnin summa. CALCULATE-funktio tarvitsee laskettavan lausekkeen ja siihen on mahdollista lisätä suodatin. Esimerkissä on käytetty SUM-funktiota, joka laskee annetun taulukon sarakkeen arvot yhteen. (CALCULATE 2022; SUM 2022.)

```
TotalSales = CALCULATE(SUM(Sales[Revenue]))
```

KUVA 6. DAX-funktio myynnin laskemiseen.

3.2.6 Visualisointien luominen

Visualisoinnit-ruudusta pystyy lisäämään haluamansa visualisoinnin. Niitä voi laittaa sivulle yhden tai useamman. Kuvassa 7 esitetään Power BI:n visualisointivaihtoehdot. Käyttäjä pystyy myös lataamaan yhteisön luomia visualisointeja käytettäväksi omaan raporttiinsa. (Visualisoinnit Power BI -raporteissa 2023.)



KUVA 7. Power BI:n visualisointivalikko.

4 DATAN VISUALISOINTI

4.1 Datan visualisoinnin hyödyt yleisesti

Ihmiset käyttävät visuaalista hahmotuskykyä jokapäiväisessä elämässä. Esimerkiksi liikenteessä ihmiset analysoivat muuta liikennettä sekä liikennemerkkejä päästäkseen haluttuun määränpäähän. Tällaisissa tilanteissa ihmiset käyttävät visuaalisia tietojenkäsittelytaitoja lähes automaattisesti. Jos sama analysointi tulisi tehdä pelkkien numeroiden ja laskutoimitusten avulla, kestäisi siinä paljon kauemmin. Se voisi myös johtaa väärään päätökseen todennäköisemmin. Tästä syystä numeroiden ja tekstin muuttaminen visualisoinneiksi auttaa ymmärtämään dataa nopeammin. (Rodriguez & Kaczmarek 2016.)

Visualisointien avulla datasta pystyy helposti erottamaan joukkoja ja toistuvuuksia. Ne voivat olla huomioitavia poikkeuksia. Esimerkiksi datan osa voi poiketa muusta datasta huomattavasti, ja se tulee yleensä ottaa huomioon. (Rodriguez & Kaczmarek 2016.)

Samana visualisoidun datan pystyy ymmärtämään useampi ihminen ympäri maailmaa. Rodriguez ja Kaczmarek (2016) toteavat, että samalla tavalla kuin matematiikka, kuvat eivät ole olennaisia vain tietyille kohderyhmälle maailmassa. Datan visualisointi on universaali tapa jakaa tietoa nopeasti, johon ei välttämättä tarvita yhteistä kieltä. (Rodriguez & Kaczmarek 2016.)

Datan visualisoinnilla voi tiivistää suuren määrän dataa yhteen kaavioon. Samalla datan useammasta eri lähteestä voi yhdistää yhteen paikkaan. Data, jonka sisältö alustavasti on monessa eri taulukossa, voidaan esittää yksinkertaisemmassa ja tehokkaammassa muodossa visualisoimalla se. (Rodriguez & Kaczmarek 2016.)

4.2 Datalla voi kertoa tarinan Power BI:ssa

Jos data ei kerro tarinaa, sen ymmärtäminen on vaikeampaa. Käyttämällä dataa voi luoda siitä kertovan tarinan. Se auttaa näyttämään datan merkityksen. (Pearson, Knight B., Knight D., Quintana 2020.)

Useassa visualisoinnissa on mahdollista näyttää työkaluvihje. Sen avulla voi tuoda lisää tietoa kohteena olevasta datasta. Vihjeen saa näkyviin viemällä hiiren kursorin esimerkiksi kaavion päälle (kuva 8). Kuten kuvasta 8 näkyy, saadaan näkyviin tarkka luku, jota ei pylväskaaviosta suoraan näe. Melkein mitä tahansa visualisointia voi käyttää vihjeenä Power BI:ssa. (Pearson ym. 2020.)



KUVA 8. Vihjeteksti pylväskaaviossa.

Kirjanmerkeillä pystyy näyttämään erilaisia versioita raporttisivusta. Raportin tekijä pystyy tallentamaan kirjanmerkkiin erilaisia asetuksia. Ne voivat olla esimerkiksi suodattimia, järjestyksiä tai näkyvyysasetuksia. Ideana on saada halutut tiedot nopeasti ja vaivattomasti esille. Kun kirjanmerkkiin on tallennettu halutut asetukset, ne saa otettua käyttöön raporttiin valitsemalla tallennetun kirjanmerkin. Power BI:n selainversiossa käyttäjät pystyvät itse tallentamaan haluamansa asetukset kirjanmerkkiin. Asetusten tallentaminen on hyödyllistä varsinkin sellaisissa raporteissa, joilla on useita käyttäjiä, jotka käyttävät sitä eri tavoin. (Pearson ym. 2020.)

4.3 Parhaat käytännöt

Power BI:lla raporttia tehdessä tulee kiinnittää huomiota visualisointien asetteluuun. Raportin tulisi olla visuaalisesti mielenkiintoinen ja toimiva samaan aikaan. Usein merkityksellisintä on, että tärkeät tiedot erottuvat ja ne ovat helposti ymmärrettävissä. Tämä on helppo toteuttaa esimerkiksi näyttämällä tärkeimmät tiedot suurimmassa koossa. Tällöin se varmasti kiinnittää käyttäjän huomion. (Vinkkejä laadukkaiden Power BI -koontinäyttöjen suunnitteluun 2023.)

Visualisointien tulee olla helppoja lukea ja ymmärtää. Siksi on tärkeä tietää, mikälaista visualisointia tulisi käyttää. Joskus yksinkertainen grafiikka, kuten esimerkiksi pylväskaavio, riittää. Erilaiset visualisoinnit sopivat käytettäväksi erilaisissa tilanteissa. (Vinkkejä laadukkaiden Power BI -koontinäyttöjen suunnitteluun 2023.)

Pylväs- ja palkkikaaviot sopivat hyvin, kun verrataan montaa eri kategoriaa. Kategorioiden arvot ovat palkkeina vierekkäin, joten niitä on helppo verrata keskenään. Pystysuuntaisesta pylväskaaviosta on helppo ymmärtää kategorian arvo verrattuna muihin kategorioihin. (Bateh & Wachsmuth 2016; Vinkkejä laadukkaiden Power BI -koontinäyttöjen suunnitteluun 2023.)

Pyöreän muotoisten kaavioiden käyttö on harvoin järkevää. Esimerkiksi ympyräkaaviota ei tule käyttää sen osien keskenään vertailuun. Sen tarkoituksena on verrata lohkon osuutta kokonaisuuteen. Sitä tulisi käyttää myös vain silloin, kun käyttää alle kahdeksaa luokkaa. (Vinkkejä laadukkaiden Power BI -koontinäyttöjen suunnitteluun 2023.)

On olemassa paljon muitakin pieniä tapoja parantaa raportin käytettävyyttä. Luvuissa tulisi käyttää enintään kolmea tai neljää numeroa. Ihmisten on helpompi lukea luku, jossa suuri määrä numeroita on korvattu tekstillä, kuten esimerkiksi merkintä ”3,4 miljoonaa”. Myös turhat arvopisteiden otsikot hämmentävät käyttäjiä. Esimerkiksi pylväsdiagrammissa käyttäjä pystyy ymmärtämään arvot ilman lukujen mainintaa. (Vinkkejä laadukkaiden Power BI -koontinäyttöjen suunnitteluun 2023.)

5 ENNEN RAPORTIN TOTEUTUSTA

5.1 Taustatietoa työstä

Teen opinnäytetyöni Sandvik Mining and Construction Oy:lle. Sandvik valmistaa laitteita kaivos- ja urakointiteollisuuteen. Näitä laitteita ovat muun muassa erilaiset kalliopora- (kuva 9), murskaus-, kuljetus- ja lastauslaitteet. Konsernin muita osaamisalueita ovat myös automaatio, digitalisaatio ja akkuteknologia. (Osaamisalueet ja innovointi n.d.)



KUVA 9. Leopard DI650i -down-the-hole porauslaite. (Sandvik 2018.)

Liiketoiminta-alueet on jaettu erilaisiin divisiooniin, jotka vastaavat liiketoiminnan eri toiminnoista. Työskentelen itse Parts & Services -divisioonassa. Se nimensä mukaisesti keskittyy varaosien ja palveluiden tuottamiseen asiakkaille ympäri maailmaa. (Parts & Services About us 2021.)

5.2 Palvelun vaihtuminen uuteen

Aikaisemmin Sandvikilla BI-työkaluna käytettiin Qlik:iä. Sillä luotiin analyttisiä raportteja myynnistä. Raporttien tarkoitus oli auttaa havainnoimaan, ymmärtämään ja vastaamaan liiketoimintaympäristöön. Se auttoi myös päätöksenteossa ja myynnin suunnittelussa. (Overview Sales Analytics QlikView Reporting n.d.)

Tiedot ohjelmistoon tulivat globaalista tietovarastosta. Käyttäjä loi taulukon datasta, joka sopi parhaiten hänen omaan käyttötarkoitukseensa. Sovelluksessa valittiin tiedot, jotka haluttiin raportissa näyttää. Esitettävän taulukon sisältöä pystyi myös rajaamaan suodattimilla. Raportti oli taulukko, joka sisälsi käyttäjän valitseman datan. Esimerkiksi jos käyttäjä valitsi tiedoiksi liiketoimintayksikön ja myynnin, muodostui taulukko näillä sarakkeilla. Tällä tavalla sai esiin kaikki liiketoimintayksiköt ja niiden myynnit. Taulukon pystyi järjestämään sarakkeiden arvojen mukaan esimerkiksi suurimmasta pienimpään. (Overview Sales Analytics QlikView Reporting n.d.)

Raportit tuli luoda aina uudelleen jokaisella käyttökerralla. Datan hakemisessa tietokannasta saattoi kestää useita minutteja. Vaikka QlikViewin raportti oli välilehdellä auki, jos sitä ei käyttänyt yli viiteen minuuttiin, sivu suoritti aikakatkaisun ja taulukko piti generoida uudelleen.

Qlikin palvelu vaihtui Power BI:hin. Sitä oli käytetty jo aikaisemminkin, mutta se oli vaatinut käyttäjän osaavan luoda itse raportit. Power BI:lla luotiin valmis itsepalvelumallin raportointipohja, jota työntekijät voivat käyttää myyntiin liittyvän tiedon tarkasteluun. Palvelua kutsutaan nimellä Business Insights. Se toimii samalla periaatteella, kuin Qlik, eli valitaan halutut tietokentät ja ohjelma luo taulukon ja kaaviot perustuen dataan. Power BI -versioon siirryttiin, koska siinä on enemmän visualisointivaihtoehtoja ja se on helpompikäyttöinen. Sandvikin Business Insightsin tietojoukkoa pystyy käyttämään lähteenä datalle omassa Power BI -raportissa. Tällöin käyttäjä pystyy luomaan itselleen kokonaisen raportin mukailtuna omaan käyttötarkoitukseensa. Tästä luvusta eteenpäin käytän Sandvikin Business Insightsista lyhennettä SBI, jotta se ei sekoitu BI-lyhenteeseen. (DATA & ANALYTICS BUSINESS INSIGHTS: MOVING TO POWER BI 2023.)

5.3 Mikä ongelma ratkaistaan?

SBI-palvelun dataa pystytään käyttämään esimerkiksi laskutusten ja kannattavuuden hinnoittelun analysointiin. Tätä dataa voidaan hyödyntää esimerkiksi verrattaessa eri myyntialueiden tai tuotteiden välisiä eroja. Myynnin kehitystä pystytään myös tarkastelemaan ja siten havainnoimaan, ovatko luvut menneet parempaan vai huonompaan suuntaan. (DATA & ANALYTICS BUSINESS INSIGHTS: MOVING TO POWER BI 2023.)

Sandvikilla on myös olemassa Power BI -työkalu Fleet Analytics. Se tuo datan Sandvikin valmistamista laitteista analytiikkapalveluun. Saatavilla on tietoa muun muassa siitä, mitä laitteita kullakin asiakkaalla on ja mikä niiden tila on. Kun datasta esimerkiksi nähdään, että laite on kulunut, pystytään sen perusteella myymään asiakkaalle muun muassa huoltopalveluita ja varaosia. (Fleet Analytics 2.0: Upcoming launch 2023.)

Tavoitteena on luoda Power BI -raportti, joka yhdistää nämä kaksi tietolähdettä. Se perustuu jo olemassa olevaan raporttiin, joka on luotu käyttäen Excel-taulukoita. Uuden raportin avulla pystyttäisiin laskemaan ja vertaamaan varaosien myyntiä suhteessa asiakkaan olemassa oleviin laitteisiin reaaliajassa. Raporttiin tulisi saada esiin asiakkaan kiviporien määrä ja kuinka monta niiden osaa on myyty edellisen 12 kuukauden aikana. Näiden tietojen pohjalta on myös tarkoitus luoda DAX-funktio, joka laskee asiakkaille myytyjen kiviporien osat suhteessa jo omistettuihin poriin. Sen avulla pystyttäisiin myös laskemaan maan sisäinen ja globaali keskiarvo tälle arvolle.

Raportti esimerkiksi auttaisi havaitsemaan, että asiakas ei ole ostanut osia yhtä paljon kuin muut asiakkaat suhteessa jo valmiiksi omistamiin laitteisiinsa. Se voi kertoa siitä, että asiakas ostaa osat muualta. Raportin tavoitteena on auttaa myynnin lisäämisessä.

6 RAPORTIN TOTEUTUS

6.1 Datan lisääminen

Datalähteinä raporttiin käytiin SBI-tietojoukkoa ja Excel-taulukoita, jotka oli tuotu Fleet Analytics -palvelusta. Näitä Excel-taulukoita oli käytetty aikaisemmin raportissa, johon työstämäni raportti perustuu. Asiakkaiden omistamien laitteiden määrä ei muutu usein, joten se ei haitannut, että sen tiedot tulivat aluksi manuaalisesti päivitetävästä Excel-tiedostosta. Tarkoituksena oli saada datamalli toimimaan ensin Excel-taulukoita käyttäen, koska se oli yksinkertaisempaa. Sen toimiessa yhteyden olisi voinut toteuttaa Fleet Analytics -palvelua käyttäen.

SBI-tietojoukko sisälsi tiedon laskutuksista, asiakkaista ja tuotteista. Tämä tietojoukko oli valmiiksi luotu, eikä sitä pystynyt muokkaamaan. Tiedot oli myös sallattu, joten en pystynyt näkemään, mitä tietojoukon taulukot sisälsivät. Tämä tietojoukko yhdistettiin raporttiin Live Connection -menetelmällä eli tieto tuli raportin ulkopuolelta. Raportissa käytetyt tiedot tältä osin olivat paikkansa pitävät. Excelistä raporttiin toin taulukot kiviporista, asiakkaista ja heidän omistamiansa laitteista. Tämä data tuotiin raporttiin Import menetelmällä, joten se tallentui raportin PBIX-tiedostoon.

Jotta nämä kaksi tietolähdettä pystyi yhdistämään, tuli niiden datamallien olla samassa paikassa. Tämän takia yhteyden muodostus SBI-tietojoukkoon vaihtui Direct Query -malliin. Tällöin taulukoiden data tuli PBIX-tiedoston ulkopuolelta, mutta datamalli oli tallennettuna tiedostoon. Silloin kaikki käytettävät taulukot olivat samassa paikassa ja SBI-taulukoiden ja Excel-taulukoiden välille pystyttiin luomaan yhteys.

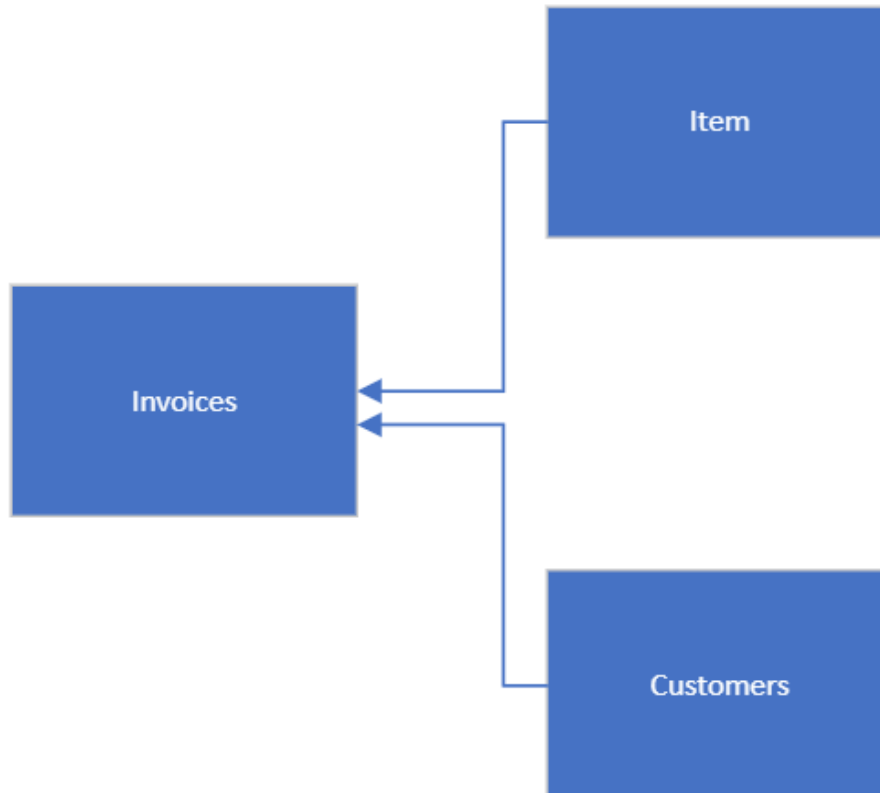
6.2 Taulukoiden väliset yhteydet

Kun tiedot oli lisätty raporttiin, seuraavaksi tuli luoda yhteydet taulukoiden välille. SBI-datasta saatujen taulukoiden yhteydet oli jo valmiiksi muodostettu, koska se

oli jo ennalta määritetty tietojoukko. Kuviossa 1 näkyy malli näiden taulukoiden yhteyksistä.

Item-taulukko sisälsi tiedot osista, Customers asiakkaista ja Invoices laskutuksista. Item ja Customers olivat molemmat yhteydessä Invoices-tauluun many-to-one -yhteydellä. Käytännössä se tarkoitti, että jokaista osaa vastasi yksi rivi Item taulukossa ja jokaista asiakasta yksi rivi Customers-taulukossa. Niiden kuitenkin oli mahdollista esiintyä useammassa eri laskutuksessa Invoices-taulukossa. Kukin Invoices-taulussa esiintyvä laskutustietojoukko puolestaan saattoi liittyä useaan eri osaan ja asiakkaaseen.

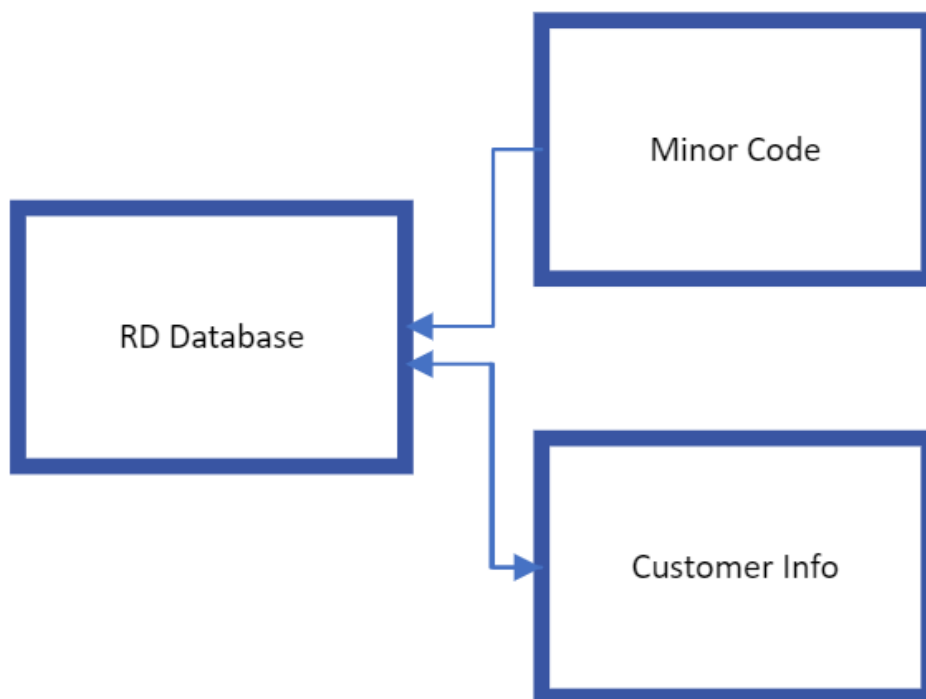
Kuvioissa 1 ja 2 nuolten suunta ja määrä kertoo taulukoiden välisen yhteyden mallin. Jos laatikoiden välisessä viivassa on yksi nuoli, yhteys on silloin many-to-one-tyyppinen. Laatikko, jota nuoli osoittaa, on näissä kuvioissa mallin many-puoli. Jos laatikoiden välisen viivan molemmissa päissä on nuoli, on yhteys silloin many-to-many, kuten kuviossa 2 näkyy.



KUVIO 1. SBI-datan taulukoiden yhteydet.

Item- ja Invoices-taulukot oli yhdistetty Item_link-nimisellä tietokentällä. Customers ja Invoices oli yhdistetty Customer_link nimisellä tietokentällä. Koska tietojoukko oli ennalta määritetty, en tiennyt mihin nämä taulukoiden yhteydet perustuivat. Samasta syystä minulla ei ollut mahdollisuutta nähdä taulukoiden sisältöä datanäkymässä. Sain taulukoiden tiedot kuitenkin näkyviin valitsemalla halutun tietokentän visualisointia tehdessäni. Kun valitsin esimerkiksi Customer_link- ja asiakkaan nimi -tietokentät, jokaisen asiakkaan nimeä vastasi joukko merkkejä. Tästä päätin, että SBI-tietojoukkoon on asiakkaille ja tuotteille generoitu tunnistetunniste, jolla taulukoiden välille pystytään helposti luomaan yhteys.

Kuviossa 2 näkyy Excel-tiedostosta lisättyjen taulukoiden väliset yhteydet. RD Database oli kiviporian tietokanta ja Customer Info sisälsi asiakkaiden tiedot ja niiden omistamat laitteet. Lisäksi Minor Code -taulukko sisälsi minor coden eli uniikin tunnisteen ja sitä vastaavan kiviporan.



KUVIO 2. Excel-taulukoiden yhteys.

Molempien tietolähteiden taulukoiden väliset yhteydet toimivat hyvin. Seuraavaksi tavoitteena oli yhdistää nämä kaksi erillistä tietolähdettä yhdeksi kokonaisuudeksi.

6.3 Tietolähteiden väliset ongelmat

6.3.1 Taulukoiden linkittäminen

Oli aika löytää ja luoda yhteys tietolähteiden välille. Jotta yhteyden pystyi luomaan, tuli myös ymmärtää, mitä tiedot taulukoiden sisällä tarkoittivat. Haasteenani oli, että en tiennyt tarpeeksi, mistä käytettävässä datassa oli kyse. Tämän ongelman ratkaisemiseen minulla oli apunani tiimi, joka tiesi, mitä taulukoissa oleva data oli. Minä taas ymmärsin, miten Power BI toimii.

Tässä vaiheessa pyrin ymmärtämään kiviporiin liittyviä tietoja ja miten ne oli tallioitu taulukoihin. Niihin liittyi esimerkiksi erilaisia mallimerkintään liittyviä kirjain- ja numeroyhdistelmiä. Täytyi myös ymmärtää yleisellä tasolla, missä kohtaa hierarkiaa käytetty tieto sijaitti.

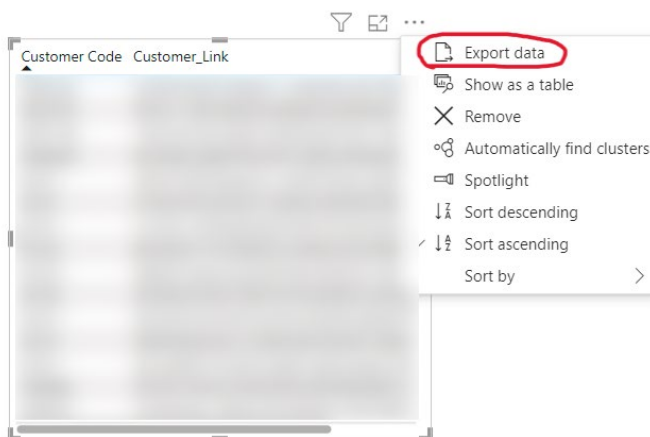
Datan ymmärtämisen tärkeys korostui, kun yritin löytää SBI-tietojoukosta Excel-tilukkoita vastaavia tietokenttiä. Koska en nähnyt tämän valmiin tietojoukon tietokenttien sisältöä muuten kuin lisäämällä sen visualisointiin, yksi tapa oli vain yrittää päätellä niitä otsikoinnin perusteella. Tämä oli työni yksi suurimmista esteistä.

Tietolähteiden väliltä löytyi yhtenäisiä tietoja. Esimerkiksi molemmissa lähteissä oli taulukko, joka sisälsi asiakkaiden tiedot. Se vaikutti järkevältä yhteydeltä, mutta ei toiminut, koska samaan aikaan data oli yhteydessä myös toisiin taulukoihin. Tällöin taulukoiden välisistä suhteista muodostui kehää muistuttava rakenne, jossa aina edellinen taulu suodatti seuraavaa. Tämänkaltaisessa tilanteessa Power BI huomauttaa virheestä ja ei anna luoda yhteyttä, joka sulkee taulukoiden muodostaman kehän. Virheilmoituksen perusteella päättelin, että tämä ei ollut oikea tapa luoda yhteyksiä näiden tietolähteiden välille.

6.3.2 Ongelmanratkointia yhteyden löytämiseksi

Vaikka SBI-tietojoukon data päivittyi, se ei kuitenkaan ollut täydellistä. Sen Item- taulukko sisälsi minor code -tietokentän, joka löytyi myös Excel-datasta, ja se vaikutti hyvältä kohdalta yhdistää tietolähteet. Tämä tietokenttä oli kuitenkin monella Item- taulukon tietorivillä tyhjä. Tätä tietoa ei ollut päivitetty tietokantaan, josta tietojoukon data tulee. Tällöin data ei ollut ehjää ja sitä ei pystytty hyödyntämään niin hyvin kuin olisi ollut mahdollista. Se pakotti löytämään uuden ratkaisun.

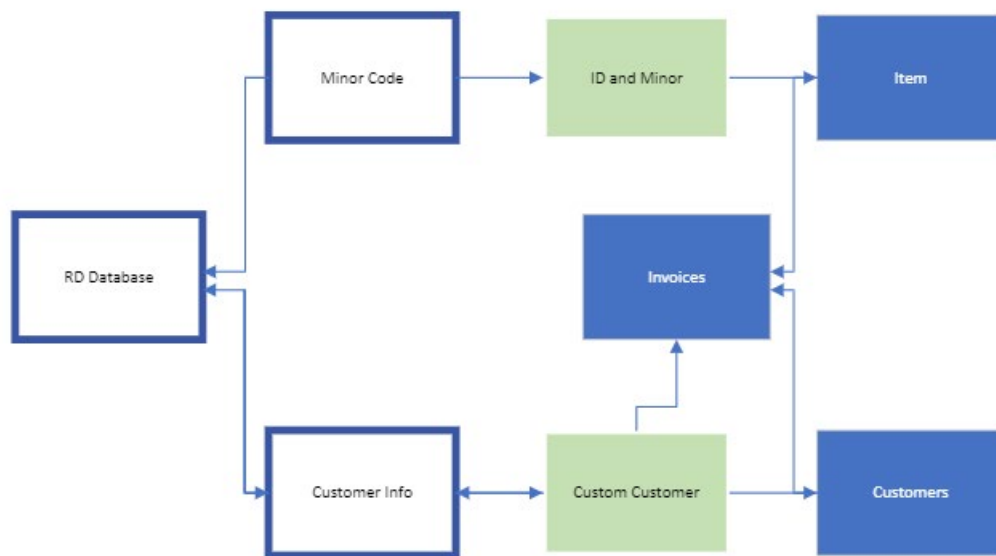
Molempien tietolähteiden asiakastauluissa oli jokaiselle asiakkaalle määritetty tunniste nimellä Customer Code. SBI:n asiakastaulu sisälsi myös Customer_link-tietokentän, josta jo mainitsin luvussa 6.2. Loin taulukkomuotoisen visualisoinnin (kuva 10), jossa oli Customer Code ja Customer_link -tietokentät. Totesin, että ne vastasivat samaa asiakasta jokaisella rivillä. Kuvassa 10 näkyvällä export data -painikkeella vein visualisoinnin taulukon csv-muotoon, jotta pystyin käyttämään sitä taulukkona, jonka sisällön näen. Latasin kyseisen tiedoston raporttiin uudeksi taulukoksi datamalliin. Tämän taulukon avulla pystyin yhdistämään Excel-tiedostoista tuodun asiakastaulukon sekä SBI:n asiakas- ja laskutustaulukot. Excel- taulukko yhdistyi tähän uuteen taulukkoon Customer Code tietokenttään ja SBI:n taulukot Customer_link tietokenttään. Tämä ei ollut paras tapa luoda yhteyttä näiden taulukoiden välille, mutta se antoi suuntaa toiminnasta. Koska Excel-tiedoston asiakastaulukko suodatti SBI:n asiakas- ja laskutustaulukkoa, kiviporien ja asiakkaiden tiedot täsmäsivät.



KUVA 10. Taulukon muuntaminen tiedostoksi.

6.3.3 Lopullinen datamalli

Loin samalla periaatteella uuden taulukon, jotta pystyin yhdistämään Minor Code ja Item -taulukot. Uusi taulukko sisälsi Minor Coden saman nimisestä taulukosta ja Item Coden Item-taulukosta. Näin sain yhteyden näiden kahden taulukon välille. Datamalli (kuvio 3) oli tämän vaiheen jälkeen jokseenkin yhtenäinen. Kuvio 3 havainnollistaa lopullista datamallia. Kuten kuvioissa 1 ja 2, laatikkoa osoittava nuoli on yhteyden many-puoli. Valkoiset laatikot ovat Excelistä tuotuja ja siniset laatikot SBI-tietojoukosta tuotuja taulukoita. Vihreät laatikot ovat luvuissa 6.3.2 ja 6.3.3 luodut taulukot.



KUVIO 3. Raportin lopullinen datamalli.

Datamalli vaikutti olevan kunnossa. Visualisointia tehdessä se näytti asiakkaan omistamien kiviporien ja laskutettujen tuotteiden määrät paikkansa pitävinä. Sen jälkeen muodostin mallia mukailevan taulukon valitsemalla halutut tietokentät. Loin myös DAX-funktiot laskemaan, kuinka monta kiviporan osaa asiakkaalta on laskutettu 12 kuukauden aikana suhteessa jo valmiiksi omistettuihin poriin. Toinen funktio (kuva 11) on samanlainen, mutta se laskee koko maan keskiarvon.

```

1 .Invoiced Quantity per Rock Drill AVG by Country =
2     CALCULATE(
3         AVERAGEX(
4             KEEPFILTERS(VALUES('Customer Info'[Customer Country])),
5             CALCULATE([.Invoiced Quantity per Rock Drill])
6         ), REMOVEFILTERS('Customer Info'[Customer Name]))

```

KUVA 11. DAX-funktio maan keskiarvosta.

6.4 Lopputulos

Sain Power BI -raportin taulukkovisualisoinnin (kuva 12) näyttämään oikeita lukuja siten, että valittuina kenttinä olivat asiakkaan maa ja nimi, minor code ja kiviporan malli. Omistettu määrä ja laskutettu määrä viimeisen 12 kuukauden ajalta ovat myös mukana. Omistettu määrä on taulukossa merkitty qty nimellä ja 12 edellisen kuukauden laskutetut tuotteet on merkitty nimellä Invoiced Quantity. Invoiced Quantity -funktio oli valmiiksi luotu SBI-tietojoukkoon. Se laskee, kuinka monta tuotetta on laskutettu riviä kohden. Kuvan 12 taulukossa tämä Invoiced Quantity -funktio laskee, kuinka monta tietyn mallin kiviporan osaa on laskutettu rivin asiakasta kohden edellisen 12 kuukauden aikana. Lisäksi taulukossa oli myös luvussa 6.3.3 itseluodut DAX-funktiot.

Customer Country	Customer Name	Item Group Minor	Model	qty	Invoiced Quantity	Invoiced Quantity per Rock Drill	Invoiced Quantity per Rock Drill AVG by Country
AUSTRALIA				6	2	0,33	0,18
AUSTRALIA				10	2	0,20	0,00
AUSTRALIA				39	1	0,03	0,00
AUSTRALIA				3	1	0,33	0,01
AUSTRALIA				3	0	0,00	0,01
AUSTRALIA				4	0	0,00	0,00
Total				973	6	0,01	0,01

KUVA 12. Taulukkovisualisointi raportissa.

Tästä kuvan 12 taulukosta jäi puuttumaan kiviporan mallin alakategoria. Se olisi kertonut tarkemmin, minkä osan asiakas on ostanut. Kuvan 12 taulukko näyttää vain minkä kiviporan osa on myyty, mutta se ei kerro mikä osa on kyseessä.

Alakategorian lisääminen lopulliseen taulukkoon ei onnistunut, koska kiviporien tietokantataulukon ja Item-taulukon välistä yhteyttä ei pystynyt toteuttamaan. Se

todennäköisesti johtui siitä, että niiden välillä ei ollut kunnollista tietokenttää, jolla sen olisi voinut luoda.

Myös toinen visualisointi jäi toteuttamatta, koska se vaati mallin alakategorian toimiakseen kunnolla. Sen ideana oli näyttää samat tiedot kuin taulukossa, mutta vain enemmän visuaalisesti. Yksinkertaisuudessaan sillä olisi voinut näyttää, kuinka suuret ovat asiakkaan ostot suhteessa omistettuihin laitteisiin per tuotekategoria. Jos asiakas olisi ollut ostanut suhteessa hyvin, olisivat palkit näkyneet vihreinä. Päinvastoin huonosti ostaneen asiakkaan arvot olisivat näkyneet punaisina.

7 POHDINTA

Yksi suurimmista haasteista oli se, että en päässyt käsiksi kunnolla dataan. Kuten aikaisemmin mainitsinkin, en tiennyt, mitä valmiin tietojoukon taulukot sisälsivät. Niiden väliset yhteydet oli tehty generoimalla tunnisteet, mikä todennäköisesti vaikeutti niiden yhdistämistä muihin datalähteisiin. Oletan, että nämä tunnisteet oli varta vasten generoitu taulukoiden välisiin yhteyksiin. Jos olisin käyttänyt taulukoita suoraan tietokannasta, olisi työ ollut varmasti helpompi.

Sen lisäksi, etten päässyt osaan datasta käsiksi, vaikeutti datan ymmärtäminen työtä hieman. Vaikka taulukon sisällön näki, jos sitä ei ymmärtänyt, ei sitä pystynyt hyödyntämään. Minulla oli apua tähän ongelmaan ja opin ymmärtämään sitä tehdessäni työtä. Vaikka opin ajan myötä ymmärtämään käyttämäni dataa, on mahdollista, että silti oli olemassa dataa, jota en tiennyt voivani käyttää hyödyksi tässä työssä.

Jos aloittaisin rakentamaan yhteyksiä taulukoiden välille uudestaan, tekisin ne ehkä hieman eri tavalla. Esimerkiksi Excelistä tuotu asiakastaulukko sisälsi asiakkaan tiedot ja omistamat laitteet. Koska oli olemassa jo toinen taulukko, joka sisälsi uniikit asiakkaiden tiedot, olisi laitteiden tiedot voinut yhdistää siihen. Toisaalta ilman Excelin asiakastaulukkoa ei olisi ollut yhteyttä kiviporien tietokantatauluun. Myös laitteille olisi voinut luoda oman taulukon, joka sisältäisi uniikkeja rivejä, koska niillä on olemassa sarjanumerot.

Vaikka kaikki ei mennyt täysin suunnitelmien mukaan, koen, että lopputulos oli oikeassa suunnassa. Se ei toimi sellaisella tavalla kuin oli tarkoitus, mutta sitä pystyy kehittämään eteenpäin. Raportista jäi tämän myötä myös visualisoinnit ja tarinankerronta vähäiseksi. Koska työ keskittyi enemmän datan mallintamiseen, olisin toivonut pääseväni hyödyntämään visualisointia enemmän. Työ kuitenkin tarjosi sopivassa määrin haastetta, ja sen kautta pystyi oppimaan vähintäänkin ongelmanratkaisutaitoja.

LÄHTEET

Aspin, A. 2022. Pro Power BI Dashboard Creation: Building Elegant and Interactive Dashboards with Visually Arresting Analytics. E-kirja. Berkeley: Apress. Viitattu 30.6.2023. Vaatii käyttöoikeuden. <https://learning.oreilly.com/library/view/pro-power-bi/9781484282274/?sso link=yes&sso link from=tampere-university>

Bateh, J. & Wachsmuth, B. G. 2016. Using statistics for better business decisions. E-kirja. New York: Business Expert Press. Viitattu 4.8.2023. Vaatii käyttöoikeuden. <https://ebookcentral.proquest.com/lib/tampere/detail.action?docID=4201910&pq-origsite=primo>

CALCULATE. 2022. Microsoft. Verkkosivu. Viitattu 5.7.2023. <https://learn.microsoft.com/en-us/dax/calculate-function-dax>

DATA & ANALYTICS BUSINESS INSIGHTS: MOVING TO POWER BI. 2023. Sandvik. PowerPoint-esitys. Viitattu 19.10.2023.

Definition decision support system (DSS). 2021. TechTarget. Verkkosivu. Viitattu 28.6.2023. <https://www.techtarget.com/searchcio/definition/decision-support-system>

Fleet Analytics 2.0: Upcoming launch. 2023. Sandvik intranet. Viitattu 19.10.2023. Vaatii käyttöoikeuden. <https://intranet.sandvik.com/aboutus/news-events/newsportal/Pages/fleet-analytics-2.0--upcoming-launch.aspx>

Foote, K.D. 2023. A Brief History of Business Intelligence. Verkkosivu. Viitattu 28.6.2023. <https://www.dataversity.net/brief-history-business-intelligence/>

Larson, B. 2020. Data Analysis with Microsoft Power BI. E-kirja. New York: McGraw-Hill Education. Viitattu 3.7.2023. Vaatii käyttöoikeuden. <https://learning.oreilly.com/library/view/data-analysis-with/9781260458626/?sso link=yes&sso link from=tampere-university>

Mikä on Power BI? n.d. Microsoft. Verkkosivu. Viitattu 29.6.2023. <https://powerbi.microsoft.com/fi-fi/what-is-power-bi/>

Mitä ovat liiketoimintatiedot? n.d. Microsoft. Verkkosivu. Viitattu 27.6.2023. <https://powerbi.microsoft.com/fi-fi/what-is-business-intelligence/>

Osaamisalueet ja innovointi. n.d. Sandvik. Verkkosivu. Viitattu 16.10.2023. <https://www.home.sandvik/fi/tietoja-meist%C3%A4/osaamisalueet-ja-innovointi/>

OVERVIEW SALES ANALYTICS QLIKVIEW REPORTING. n.d. Sandvik. PowerPoint-esitys. Viitattu 18.10.2023.

Parts & Services About us. 2021. Sandvik intranet. Viitattu 16.10.2023. Vaatii käyttöoikeuden. <https://intranet.sandvik.com/org-info/smrt/pa/parts-services/about-us/Pages/default.aspx>

Pearson, M., Knight, B., Knight, D. & Quintana, M. 2020. Data Storytelling with Power BI. E-kirja. Berkeley: Apress. Viitattu 25.7.2023. Vaatii käyttöoikeuden. https://learning.oreilly.com/library/view/pro-microsoft-power/9781484260081/?sso_link=yes&sso_link_from=tampere-university

Pratt, M.K. 2019. What is business intelligence? Turning data into business insights. Cio. Verkkosivu. Viitattu 29.6.2023. Vaatii käyttöoikeuden. <https://www.proquest.com/comps-cijour/docview/2306008661/360A498712CD4E6CPQ/1?accountid=14242>

Qlik Sense® | Modern Cloud Analytics. n.d. Qlik. Verkkosivu. Viitattu 29.6.2023. <https://www.qlik.com/us/products/qlik-sense-analytics?ga-link=hp-hero>

Rodriguez, J. & Kaczmarek, P. 2016. Visualizing financial data. E-kirja. Indianapolis: Wiley. Viitattu 25.7.2023. Vaatii käyttöoikeuden. https://learning.oreilly.com/library/view/visualizing-financial-data/9781118907856/?sso_link=yes&sso_link_from=tampere-university

Sayer, P. & Olavsrud, T. 2021. Top 12 BI tools. Cio. Verkkosivu. Viitattu 29.6.2023. <https://www.cio.com/article/222558/top-business-intelligence-bi-tools.html>

Suhteiden luominen ja hallinta Power BI Desktopissa. 2023. Microsoft. Verkkosivu. Viitattu 5.7.2023. <https://learn.microsoft.com/fi-fi/power-bi/transform-model/desktop-create-and-manage-relationships>

SUM. 2022. Microsoft. Verkkosivu. Viitattu 5.7.2023. <https://learn.microsoft.com/en-us/dax/sum-function-dax>

Vinkkejä laadukkaiden Power BI -koontinäyttöjen suunnitteluun. 2023. Microsoft. Verkkosivu. Viitattu 3.8.2023. <https://learn.microsoft.com/fi-fi/power-bi/create-reports/service-dashboards-design-tips>

Visualisoinnit Power BI -raporteissa. 2023. Microsoft. Verkkosivu. Viitattu 25.7.2023. <https://learn.microsoft.com/fi-fi/power-bi/visuals/power-bi-report-visualizations>

What is business intelligence? n.d. IBM. Verkkosivu. Viitattu 27.6.2023. <https://www.ibm.com/topics/business-intelligence>

What is Tableau? 2023. Tableau. Verkkosivu. Viitattu 29.6.2023. <https://www.tableau.com/why-tableau/what-is-tableau>