

# **Valtion talousarvion avoin data seurannassa ja ennustamisessa**

Case: Suomen Pankki



Ammattikorkeakoulun opinnäytetyö

Tietojenkäsittelyn koulutusohjelma

Hämeenlinna, syksy, 2017

Anna Marin

Tietojenkäsittely, tradenomi  
Visamäki, Hämeenlinna

---

<b>Tekijä</b>	Anna Marin	<b>Vuosi</b> 2017
<b>Työn nimi</b>	Valtion talousarvion avion data ennustamisessa ja seurannassa	
<b>Työn ohjaaja/t</b>	Lasse Seppänen	

---

## TIIVISTELMÄ

Työn taustalla oli Suomen Pankin ennustetoimiston tarve kehittää talousarviodatan käyttöä ennuste- ja seurantatyössä. Kehitystyö toteutettiin avoimen talousarviodatan käyttöönottona Power BI-välineellä. Välineellä toteutettiin ennuste- ja seurantatyön tueksi visuaaliset perusraporttinäytöt ja koontinäyttö (Dashboard). Työssä arvioidaan Power BI:n käyttöä ja syntyneitä lopputuotteita (raporttinäytöt ja Dashboard).

Power BI:n vahvuudet ovat visuaalisessa esittämisessä. Lopputuloksena syntyneet raportit sekä koontinäyttö vastasivatkin hyvin seuranta- ja ennustetyön perustarpeita. Datan käsittelyn osalta Power BI vastasi osittain talousarviodatan käytölle asetettuihin vaatimuksiin. Sen tehokkaampaa käyttöä edistäisi raakadatan käsittelyn helpottaminen, laskennallisten ominaisuuksien parantaminen ja datan käsittelyn liittyvien ylläpito- ja muokkaustoimintojen kehittäminen.

Työn teoriaosassa käsitellään avoimen datan taloudellisia vaikutuksia. Teoriaosuuden pohjalta tehdään päätelmiä avoimen talousarviodatan käyttöönoton taloudellisesta kannattavuudesta.

Avoimen datan taloudellisista vaikutuksista ei ole juurikaan tutkimustuloksia. Ennakoivien arviointien mukaan avoimen datan käytöllä on mahdollista saavuttaa taloudellisia hyötyjä mm. toimintojen tehostuessa. Toimintojen tehostaminen oli myös avoimen talousarviodatan käyttöönoton yhtenä tavoitteena. Vaikka alkuvaiheessa syntyi mm. talousarviodatan käyttöönottoon liittyviä työkustannuksia, voivat datan käytöstä ja jatkojalostetusta aineistosta saatavat hyödyt maksaa itsensä takaisin hyvinkin lyhyessä ajassa.

**Avainsanat** avoin data, Power BI, talousarviodata, visualisointi.  
**Sivut** 52

Degree Programme in Business Information Technology  
Visamäki, Hämeenlinna

---

<b>Author</b>	Anna Marin	<b>Year</b> 2017
<b>Subject</b>	State Open Budget Data with Follow-up Work and Forecasting	
<b>Supervisor</b>	Lasse Seppänen	

---

ABSTRACT

Background of this bachelor's thesis was in the Bank of Finland's forecasting office's need to develop the use of public budget data. Development work was put in practice with the use of open budget data and Power BI-program. The end results were the visual basic reports and a Dashboard to help the forecasting and its follow-up work. This thesis evaluates the use of Power BI and the generated end products (visual reports and Dashboard).

Power BI's strengths are its visual presentations. The end reports as well as the Dashboard responded well to the basic needs of follow-up and forecasting work. Regarding data processing, Power BI partly met the budget data usage requirements. Its more efficient use in a wider user base would help to ease the processing of raw data, improve computational capabilities, and develop data processing. The use of the tool would also improve the maintenance and editing functions related to the data.

The theory of this thesis deals with economic effects of the open data. Based on the theoretical part, conclusions are drawn on the economic viability of using an open budget data.

There is hardly any research data on the economic impact of using open data. According to proactive estimates, by using the open data it is possible to achieve economic benefits and efficiency enhancement. Improving the efficiency of operations was also one of the goals of the use of open budget data. Although the usage of open data causes labor costs in the early stages, the benefits resulting from the usage of open data can repay those costs in a very short time.

**Keywords** Open data, Power BI, budget data, visualization.

**Pages** 52

# SISÄLLYS

1	JOHDANTO.....	1
1.1	Työmenetelmät, tutkimuskysymykset ja rajaukset .....	2
2	AVOIN DATA .....	3
2.1	Avoimen datan määritelmä.....	4
2.2	Avoimen datan lisenssi.....	5
3	AVOIMEN DATAN TALOUDELLISET VAIKUTUKSET .....	6
3.1	Avaamisesta ja ylläpidosta aiheutuvat kustannukset .....	6
3.2	Avoimen datan käytön taloudelliset vaikutukset yrityksissä .....	7
3.3	Kansantaloudelliset vaikutukset .....	8
3.4	Johtopäätökset avoimen datan taloudellisista vaikutuksista .....	10
4	KÄYTTÖÖNOTETTAVA AVOIN TALOUSARVIODATA.....	11
4.1	Valtion talousarvio .....	11
4.2	Talousarviodatan käytön lähtötilanne Suomen Pankissa .....	11
4.3	Tietotarpeiden määrittely .....	12
4.4	Kuinka hyvin avoin talousarviodata täyttää avoimen datan vaatimukset?.....	12
5	POWER BI.....	14
5.1	Business Intelligence ja Power BI .....	14
5.2	Power BI pähkinänkuoressa .....	15
5.2.1	Power BI Desktop .....	16
5.2.2	Power BI Service ja Dashboard.....	16
6	TALOUSARVIODATAN KÄYTTÖÖNOTTO POWER BI:LLA.....	18
6.1	Työvaiheet Desktopilla .....	18
6.2	Datan käyttöönoton toteutus Power BI:lla .....	20
6.3	Datan sisällön tarkastus .....	21
7	TALOUSARVIODATAN KÄSITTELY POWER BI:LLA .....	22
7.1	Talousarvion taulujen käsittely Power BI:lla .....	22
7.2	Laskenta Power BI:lla .....	24
8	TALOUSARVION PERUSRAPORTTIEN JA KOONTINÄYTÖN TOTEUTUS .....	26
8.1	Raporttinäyttöjen visualisointien toteutus Power BI:lla.....	26
8.2	Hierarkianäkymän toteutus .....	33
8.3	Kustomoidut visualisoinnit.....	35
8.3.1	Puutyypinen virtauskaavio .....	35
8.3.2	Verkostomalli.....	36
8.3.3	Yhdistelmägrafiikka .....	37
8.4	Talousarvion raporttien julkaiseminen Power BI -pilvipalvelussa .....	38
8.5	Talousarvion koontinäytön teko Power BI -pilvipalvelussa .....	39

9	KOKEMUKSET POWER BI:N KÄYTÖSTÄ.....	42
9.1	Taulujen käsittely .....	42
9.2	Laskenta.....	43
9.3	Visualisointi .....	43
9.4	Pilvipalvelu .....	44
9.5	Koontinäyttö.....	44
10	POWER BI:N HAASTEET .....	45
10.1	Datan käsittelyn ja kuviohallinnan haasteet .....	45
10.2	Pilvipalvelun haasteet .....	45
10.3	Käyttäjätuki .....	46
11	TULOKSET .....	47
11.1	Tulosten arviointi .....	47
11.2	Jatkokehityshankkeet.....	48
12	YHTEENVETO .....	49
	LÄHTEET .....	50
	KUVALÄHTEET.....	52

## KÄSITELUETTELO

Lyhenne	Lyhenteen selitys
Ad hoc	Ennalta määrittelemättömät, nopeasti tehtävät toimeksiannot.
Aikasarja	Havaintoyksiköstä tehtyjen mittausten muodostama aineisto.
Avoin data	<b>Open data</b> on maksutonta, koneluettavaa, laillista, kaikkien käytettävissä olevaa dataa.
Avoin lähdekoodi	Tiokoneohjelmien tuottamis- ja kehittämismenelmiä, jotka tarjoavat käyttäjille mahdollisuuden tutustua ohjelman lähdekoodiin ja muokata sitä omien tarpeiden mukaan. Periaatteena vapaus käyttää ohjelmaa mihin tahansa tarkoitukseen.
Avoin ohjelmointirajapinta	Ohjelmointirajapinta, jonka kaikki ominaisuudet ovat julkisia.
BI	<b>Business Intelligence</b> , liiketoimintatiedon hallinta. Laajempänä käsitteenä tiedon jalostusprosessi ja tietovarastojen rakentaminen.
CC BY 4.0	Avoimelle datalla suositeltu lisenssi. CC BY 4.0 lisenssin käyttöluvan kohteena olevaa dataa voi jakaa vapaasti ja maksutta.
Crative Commons	Maailman laajuinen standardi avoimen tiedon julkaisuun.
Dashboard	Keskeisten tunnuslukujen seurantaan tarkoitettu koontinäyttö. Sisältää mm. visuaalisia mittaristoja, kaavioita jne.
Dataformaatti	Tiedostomuoto johon dataa tallennetaan. Esimerkiksi CSV, JSON ja XML.
ETL	(Extract - Transform - Load) tietojen jalostaminen tietolähteistä tietovarastoon.
Koneäly	Synonyymi tekoälylle ja keinoälylle. Tarkempi määrittely on avoin. Eräessä merkityksessä tietokoneohjelma joka suorittaa yleisesti ihmismäistä ajattelua.
Lähdekoodi	Tietokoneohjelman lähdekoodi sisältää käskyt ja ohjeet, joiden avulla tietokone saadaan toimimaan halutulla tavalla. Lähdekoodi on useimmiten käännettävä suoritettavaan muotoon ennen sen ajamista tietokoneessa.
Massadata	<b>Big data</b> valtava datamäärä, joiden yhteydessä ei voida soveltaa perinteisiä datan hallintatapoja.

Lyhenne	Lyhenteen selitys
Metadata	Tietoa tiedosta. Metadata kuvaa tietojen sisällön ja merkityksen.
Rajapinta	Ohjelmointirajapinta (Application programming interface, API) määrittelee, miten ohjelmisto tarjoaa tietoja tai palveluita sovelluksille tai tietojärjestelmille. Laajempi kuvaus osoitteessa: <a href="http://avoinrajapinta.fi/">http://avoinrajapinta.fi/</a>
Rest	<b>Representational State Transfer</b> , hybridi-arkkitehtuurimalli, joka lainaa hyväksi todettuja eri arkkitehtuurimallien ohjelmistovaatimuksia, yhdistäen ne toimivaksi kokonaisuudeksi.
Tietotilinpäätös	Tietotilinpäätös on raportti, joka syntyy organisaation sisäisen tarkastelun tuloksena ja antaa kokonaiskuvan organisaation tietojenkäsittelyn nykytilasta.
XML	<b>Extended Markup Language</b> . Yleinen standardi jolla määritellään tapa, jolla tietoja siirretään organisaatioiden välillä.

## 1 JOHDANTO

Tämän opinnäytetyön toimeksiantaja on Suomen Pankki. Suomen Pankki on kansallinen keskuspankki ja se on eurojärjestelmän aktiivinen jäsen. Suomen Pankin Rahapolitiikkaosaston ennustetoimistossa tehdään mm. Suomen kokonaistaloudellisia ennusteita. Toimiston tehtäviin kuuluu Valtion talousarvioidatan seuraaminen, joka on osa julkisen talouden seuranta- ja ennustetyötä.

Opinnäytetyön idea syntyi tarpeesta kehittää Valtion talousarvioidatan käyttöä ennustetoimistossa. Dataa on aikaisemmin kerätty käsin ja kerätystä aineistosta ei ole saatu irti kaikkea mahdollista tietoa ennustamisen ja seurannan tueksi. Valtiovarainministeriö tarjoaa huomattavasti kattavammassa mittakaavassa avointa talousarviodataa, ja sen käyttöönottoa onkin mietitty ennustetoimistossa jo pidemmän aikaa.

Tämän opinnäytetyön tavoitteena on toteuttaa avoimen talousarvioidatan käyttöönotto ja jatkojalostaminen Power BI-analyysivälinettä hyödyntäen Suomen Pankin ennustetoimistossa. Jatkojalostetusta aineistosta rakennetaan ennuste- ja seurantatyön analyysin avuksi visuaaliset perusraporttinäytöt, joista koostetaan loppukäyttäjän seurattavaksi koontinäyttö (Dashboard). Työn painotus on toiminnallisessa opinnäytetyössä.

Työn teoriaosuudessa käsitellään avoimen datan taloudellisia vaikutuksia. Teoriaosuuden tavoitteena on tuoda esiin arvioita siitä, mitä kustannuksia avoimen datan avaamisesta syntyy ja mitä taloudellista hyötyä sen käytöllä voidaan saavuttaa. Näiden arvioiden pohjalta saadaan näkemyksiä myös avoimen talousarvioidatan käytön kannattavuudesta.



## 1.1 Työmenetelmät, tutkimuskysymykset ja rajaukset

Tutkimuksessa käytetään kvalitatiivista eli laadullista menetelmää. Teoriaosuutta käsitellään Valtiovarainministeriön ja Valtioneuvoston kanslian sekä Euroopan komission tekemien tutkimusraporttien pohjalta. Lisäksi teoriaosuudessa käytetään aihetta käsittelevää kirjallisuutta, opinnäytetöitä ja internetsivustoja sekä avoimen datan portaalreja.

Työn teoriaosuudessa pyritään vastaamaan seuraavaan tutkimuskysymykseen:

- Mitkä ovat avoimen datan taloudelliset vaikutukset?

Käytännön osuuden tavoitteena on selvittää, miten talousarviodatasta saataisiin Power BI -välinettä käyttäen hyödyllistä sekä paremmin tulokittavaa aineistoa ennuste- ja seurantatyön tueksi. Empiirisen osion aineistomateriaaleina on käytetty Microsoftin Power BI -käyttöoppaita, kirjallisuutta, opinnäytetöitä sekä haastatteluja.

Empiirisessä osiossa selvitetään Suomen Pankin ennustoimiston case-esimerkin avulla:

- Miten Power BI:n avulla voidaan hyödyntää avoimen talousarviodatan käyttöä seuranta- ja ennustetyössä?

Loppupäätelmiä uudella tavalla jalostetun ja käytetyn avoimen datan hyödyllisyydestä tehdään opinnäytetyön kirjoittajan analyysin pohjalta.

Tässä työssä käsitellään Power BI Free tuoteperhettä. Maksullisen Power BI Pron ominaisuuksien selvitys jää lyhyen käytön ja kirjallisuuden pohjalta tehtyihin päätelmiin. Pilvipalvelun julkaisualustana (Power BI Service) käytetään Hämeenlinnan ammattikorkeakoulun Office365 pilvipalvelua.

## 2 AVOIN DATA

Avoim data on elektronisessa muodossa olevaa tietoaineistoa, jota kuka tahansa voi hyödyntää maksutta lain sallimissa puitteissa. Avoimen datan ajatus syntyi 1980-luvulla. Sen taustalla olivat liikkeet, jotka edistivät avoimen lähdekoodin ja ohjelmistojen kehitystä ja käyttöä. Dataa on avannut erityisesti julkinen sektori. Osasy julkishallinnon datan avaamiselle löytyy 2000-luvun alkupuolelta. Tuolloin Yhdysvallat ja Iso-Britannia avasivat datavarantojaan edistääkseen poliittista läpinäkyvyyttä ja kansalaisyhteiskuntaa. Avoimen datan kuvaukset määriteltiin ensimmäisenä maailmassa vuonna 2004 Open Knowledgein toimesta. Tuolloin määriteltiin mm. "avoimen datan" käsitteen sisältö. (Helsinki Region Info Share, HRI 2011; A University for the Web, P2PU 2017; Euroopan komissio 2015, 25.)

Suomessa avoimen datan avaamista kiihdytti vuonna 2009 kansainvälinen kehitys avoimen datan käyttöönotossa, sekä Euroopan unionin PSI-direktiivi (julkisen sektorin tietojen uudelleenkäyttöä ohjaava EU-direktiivi). Hallitus julkistikin vuonna 2011 periaatepäätöksen julkisen sektorin digitaalisen tietoaineiston saatavuuden parantamisesta ja uudelleen käytön edistämisestä. Päätös sysäsi eteenpäin julkishallinnon avoimen datan käytön kehittämistä. Maanmittauslaitos oli julkishallinnossa datan avaamisen edelläkävijä ja se avasi ensimmäisenä maastotietojaan avoimena datana vuonna 2012. (P2PU 2017.) Suomessa julkishallinto jolla on hallussaan laajat tietovarastot on vuoteen 2017 mennessä avannut mm. maastotietoja, säätietoa, liikenne- ja ajoneuvotietoa, tilastoja, kulttuuriaineistoa, julkisia verotietoja sekä lakitietoa (VM 31/2015, 10).

Vuonna 2013 käynnistettiin Valtiovarainministeriön vetämä avoimen tiedon ohjelma 2013 – 2015. Ohjelman lähtötavoitteena on ollut avata vuosikymmenen loppuun mennessä kaikki merkittävät julkishallinnon julkiset tietovarannot, ottaen huomioon lainsäädännön rajoitukset. (VM 31/2015, 9.)

Viimeisen kymmenen vuoden aikana avoimen datan kehityskulku on ollut nopeaa, niin Suomessa kuin kansainvälisesti. Datan tarjoamiseen panostetun työn jälkeen julkishallinto on siirtynyt painottamaan avoimen datan tiedon hyödyntämiseen, tuottamiseen, jakamiseen, avaamiseen ja analysointiin liittyvien taitojen kasvattamista (VM 31/2015, 30).

Dataa on saatavissa runsaasti, eikä näin ollen tiedon määrä välttämättä ole ongelmakohta. Sen sijaan haastavaksi saatetaan kokea isojen datamäärien hallinta ja järkevä haltuunotto. Uudeksi nousevaksi ilmiöksi mm. avoimen datan käytössä onkin noussut isojen tietomassojen käyt-

töönotto, datan reaaliajassa tapahtuva analysointi ja sen visualisoiminen mahdollisimman ymmärrettävässä muodossa. Tämän vuoksi Business Intelligence -ratkaisut ovat nousseet yritysmaailmassa tärkeikeiksi ICT-investointikohteiksi. (Romppainen 2015, 9.)

## 2.1 Avoimen datan määritelmä

Avoimella datalla tarkoitetaan julkishallinnolle, organisaatioille, yrityksille tai yksityishenkilöille kertynyttä tietoa, joka on avattu tasavertaisesti kaikkien käytettäväksi. Sitä voi hyödyntää maksutta omia käyttötarkoituksia tai kaupallista käyttöä varten. (HRI 2010.) Avoin data on laillista, uudelleen käytettävää, maksutonta, koneluettavaa ja helposti löydettävää sekä vapaasti käytettävää ja jaettavaa dataa. Sen katsotaan edistävän mm. demokratiaa, hallinnon läpinäkyvyyttä ja tuottavuutta. (Pelttari 2014, 3.)

Avoimen datan tulee olla julkista tietoa, se ei saa vaarantaa kenenkään yksityisyyden suojaa tai yleistä turvallisuutta. Datassa ei saa olla henkilötietoja tai liikesalaisuuksia. Teknisen saatavuuden näkökulmasta dataa tulee pystyä käsittelemään helposti tietokoneohjelmilla. Koneelliseen hyödynnettävyyteen sopivia muotoja ovat esimerkiksi: CSV-, XLS- tai XML-muodot. Lisäksi soveltuvia ovat myös ohjelmointirajapinnat suoraan lähteeseen. Datan rakenne ja merkitys tulee kuvata ymmärrettävästi. Kuvailussa käytetään metadataa. Metadataa voivat olla esimerkiksi kirjan tiedot (kirjan nimi, kirjailija, kustantaja, julkaisu päivämäärä). Avoimen datan sijainnista ja olemassaolosta tulee kertoa yleisesti. Löydettävyys paranee kun se lisätään julkiseen katalogiin. (HRI 2010.) Avoimen datan kriteerit on esitetty tiivistetyssä muodossa kuvassa 1.



Kuva 1. Avoimen datan kriteerit (mukaillen Romppainen 2015; HRI 2010).

## 2.2 Avoimen datan lisenssi

Avoimen datan yhteydessä on hyvä käyttää lisenssiä, joka kertoo mitä avoimella datalla saa tehdä. Käyttöoikeuksien puuttuessa saattaa avoin data jäädä käyttämättä oikeuksien selvittämisen työläyden vuoksi.(HRI 2010.)

Käyttölupaa laadittaessa on suositeltavaa käyttää Creative Commons -lisenssi-perhettä, joka on maailman laajimmin käytetty avoimen lisenssin sisältöjärjestelmä. Se on myös oikeuskäytännöissä toimiva lisenssi-perhe. Lisenssivaihtoehtoja on yhteensä 7 kappaletta. Oikean lisenssin valitsemisessa auttaa Creative Commonsin internetsivulla oleva lisenssi-valitsin. Lisenssin käyttöön ei tarvita mitään lupaa tai maksua. (creativecommons.fi 2017.)

Avoimen datan osalta kannattaa ottaa huomioon se, että jos aineiston kaupallista käyttöä tai muuntelu-oikeutta rajoitetaan, ei tällöin voida puhua avoimesta datasta. Avoimen datan kohdalla tulee käyttää lisenssiä CC BY, joka tarkoittaa sitä, että dataa voi jakaa vapaasti kunhan tekijän nimi ja lisenssi mainitaan datan yhteydessä.(creativecommons.fi 2017.)

### 3 AVOIMEN DATAN TALOUDELLISET VAIKUTUKSET

Avoimen datan julkaiseminen vaatii käytännössä aikaa, rahaa ja osaamista. Tämän vuoksi tulisi avoimen datan julkaisemisen kustannuksista sekä datan käytöstä saaduista taloudellisista hyödyistä olla konkreettisia todisteita. Organisaatiot ja julkisyhteisöt, jotka suunnittelevat datan avaamista, haluavat useimmiten vastinetta työpanokselleen sekä kustannuksien kattamiseksi. Datan taloudellisia vaikutuksia on mitattava, seurattava ja arvioitava (VM 15a2015, 1, 2).

Saatavissa olevia tutkimustuloksia avoimen datan taloudellisista vaikutuksista on vähän. Siitä on tehty ennakoivaa arviointia, mutta todellisista vaikutuksista ei ole juurikaan tutkimustietoa. (VM 15a2015, 8.; VNK 2017, 22.)

Tutkimustietoa tarvitaan siitä, miten avointa dataa on käytetty ja mitä taloudellisia hyötyjä sen käytössä on saavutettu. Tämän lisäksi tulisi kerätä tietoa siitä, mitä kustannuksia datan avaamisesta ja sen ylläpidosta on syntynyt. Taloudellisten vaikutusten mittaamiseen tarvitaan välineitä ja menetelmiä. Luotettavien tulosten aikaansaamiseksi välineiden ja menetelmien tulisi olla oikein valittuja. Siksi niitä tulisi tutkia sekä kehittää. (VNK 2017, 20.) Vasta näiden valintojen jälkeen avoimen datan taloudellisista vaikutuksista on mahdollista kerätä luotettavaa tietoa. Kerätyn tiedon pohjalta voidaan tehdä johtopäätöksiä kustannustehokkaasta tietovarastojen avaamisesta ja ylläpidosta, sekä datan avaamisen kannattavuudesta. Kerätty tutkimustieto auttaa saavuttamaan parhaimmat yhteiskunnalliset ja taloudelliset hyödyt.

#### 3.1 Avaamisesta ja ylläpidosta aiheutuvat kustannukset

Datan avaamisessa kannattaa panostaa laatuun ja käytettävyyteen. Laadukkaan ja helposti käyttöön otettavan datan mahdollistamiseen tarvitaan työn toteuttajia eli osaavia resursseja. Tämä saattaa tarkoittaa käytännössä lisäresursseihin ja koulutukseen liittyviä kustannuksia. Vastapainoksi datan avauksen jälkeen dataan kohdistuvat tietopyynnöt ja kyselyt vähenevät. Tämän seurauksena neuvonnan tarve yleensä vähenee, eikä siihen tarvita resursseja yhtä paljon kuin aikaisemmin on tarvittu. (VM 31/2015, 16.)

Myös uusien tietojärjestelmien kehittäminen lisää kuluja. Datavarantojen suuret käyttäjämäärät saattavat lisätä datan suhteen uusia tietotarpeita. Käyttäjät toimivat eri tietojärjestelmäympäristöissä, joten datan siirrettävyyden osalta saattaa tulla lisävaatimuksia. Lisävaatimukset kasvattavat jakelujärjestelmien ja käytettävien ohjelmajapintojen kehittämiseen ja ylläpitoon liittyviä menoja. Avoimen datan käyttöönoton jälkeen tulee huolehtia datan laadusta, ajantasaisuudesta,

luotettavuudesta ja ylläpidosta. Tämä tarkoittaa tiedon sisältöön liittyvää työtä, jolla on oma hintansa.(VM 31/2015, 16.)

### 3.2 Avoimen datan käytön taloudelliset vaikutukset yrityksissä

Valtioneuvoston kanslian vuonna 2017 julkaisemassa raportissa "Avoimen datan hyödyntäminen ja vaikuttavuus" on tutkittu avoimen datan käytön taloudellisia vaikutuksia yritysten innovaatiotoimintaan ja kannattavuuteen. Tutkimustulokset pohjautuvat Tilastokeskuksen vuonna 2014 toteuttaman innovaatiotutkimuksen, ja Etlan vuonna 2011 pienyrityksille teettämän yrityskyselyn tuloksiin. (VNK 2017, 22-23, 27.)

Innovaationäkökulmasta katsottuna uusien palveluiden ja tuotteiden tuottaminen ei suoranaisesti vielä tuota taloudellista hyötyä yrityksille. Niiden katsotaan kuitenkin lisäävän yritysten kilpailukykyä ja voitontekomahdollisuuksia ja näin ollen myös yritysten kannattavuutta. Täysin uusia merkittäviä innovaatiota esimerkiksi tavaroiden ja palveluiden tuottamisessa keksivät organisaatiot ovat käyttäneet avointa dataa huomattavasti enemmän kuin muut yritykset. Huomattavaa on myös se, että avointa dataa omassa innovaatiotoiminnassaan hyödyntävät palvelualan yritykset ovat tehneet enemmän innovaatioita, kuin tavaran tuotantoa harjoittavat yritykset. (VNK 2017, 23-25.)

Innovaatiotutkimuksen mukaan avoimen datan suurin käyttäjäryhmä on yli 250 hengen organisaatiot ja nuoret alle 5 vuotta vanhat yritykset (VNK 2017, 23). Usein juuri nuoret yritykset aloittavat toimintansa uuden innovaation tai kehitysidean voimin. Isoilla yrityksillä on yleensä enemmän pääomaa ja resursseja käytettävissä tuotteiden ja palveluiden kehitystyöhön, joten on luonnollista, että myös avoimen datan hyödyntäminen voidaan huomioida isoissa yrityksissä paremmin. Lisäksi on huomattu että ne yritykset, jotka verkostoituvat, sähköistävät eli digitalisoivat toimintojaan tai hyödyntävät massadataa, käyttävät enemmän avointa dataa innovaatiotoiminnassaan (VNK 2017, 23).

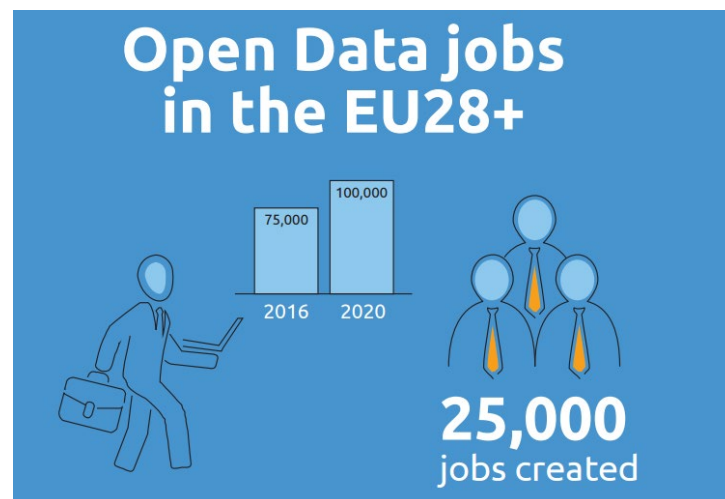
Etlan pienyrityksille teettämän kyselyn mukaan julkista dataa (mukana myös avoin data) tuote- ja palvelukehitystyössään käyttäneet yritykset eivät ole kasvaneet muita yrityksiä nopeammin. Tosin saatavissa olevaa julkisen sektorin avointa dataa, kuten *Maanmittauslaitoksen digitaalisia maastoaineistoja* ja *Ilmatieteen laitoksen säätietoja* on ollut saatavissa avoimena datana vasta vuoden 2011 jälkeen. (VNK 2017, 28, 31.)

Tuoreempaa tietoa on saatu liikennetietoja tuote- ja palveluinnovaatioissaan hyödyntävistä pienistä yrityksistä. Nämä yritykset ovat pystyneet kasvattamaan omaa liikevoittoaan huomattavasti verrattuna muihin vastaaviin pienyrityksiin. Trafín vuoden 2015 tietotilinpäätös tukee tätä tietoa avoimen datan latausten suuren määrän osalta. (VNK 2017,

30.) Tämän tiedon valossa on hyvinkin todennäköistä, että myös paikatietoja, maastoaineistoja ja säätietoja käyttävien pienten yritysten liikevaihto on kasvusuuntaista.

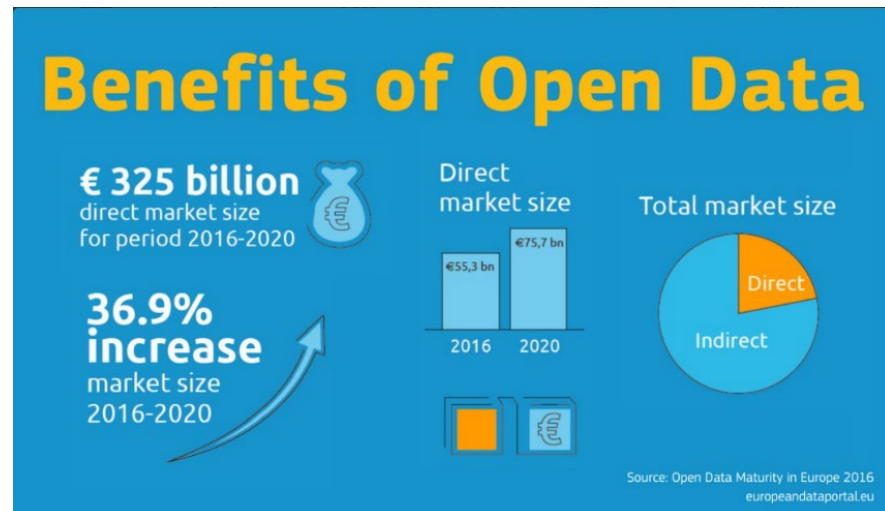
### 3.3 Kansantaloudelliset vaikutukset

Avoimen datan kansantaloudellisista vaikutuksista tuottavuuteen tai työllisyyteen ei ole kerätty juurikaan tietoa. Tämä ilmenee mm. vuonna 2013 tehdyn avoimen datan barometrista, jossa oli mukana 77 maata. Ennakollista arviointia sen sijaan on tehty sekä maailmanlaajuisesti että kansallisesti. Globaalisti avoimen datan markkinoiden arvoksi on arvioitu jopa kolme triljoonaa euroa (mukana julkisen- sekä yksityisen sektorin avoin data). (VM 15a/2015, 12–13, 20.) Työllistäviä vaikutuksia arvioidaan suoralla avoimen datan käytöllä syntyvän vuosien 2016–2020 aikana 25 000 työpaikkaa EU28+ -maissa (Kuva 2). Tulokset työllistävästä vaikutuksesta perustuvat Tanskan tekemien arvioiden pohjalta laskettuihin estimaatteihin, joten niiden luotettavuutta on kyseenalaistettava. (Euroopan komissio 2015, 78–79.)



Kuva 2. Avoimen datan työllistävät vaikutukset (avoimen datan suora käyttö) (Euroopan komissio 2015).

Avoimella datalla uskotaan olevan merkittävä taloudellinen arvo. Eurooppalaisen Data Portaalin mukaan, vuonna 2016 avoimen datan markkinoiden arvioidaan olevan 55,3 miljardin kokoluokkaa. Vuoteen 2020 mennessä kasvuksi on arvioitu 37 %. Markkinoiden koko olisi tuolloin noin 75,7 miljardia euroa (Kuva 3). (Euroopan komissio 2015, 76.)



Kuva 3. Avoimen datan markkinoiden kasvu (Euroopan komissio 2015).

Isona hankkeena voidaan pitää USA:n ja EU:n välillä olevaa yhteisten tietokantojen avaushanketta, jonka tavoitteena on luoda yhteinen avoimen lähdekoodin tuote, joka helpottaa EU:n ja USA:n tietojen avaamista ja aineistojen hakua. Euroopan Komission ja Yhdysvaltain kauppaministeriön asiantuntijat ovat kehittäneet avoimen lähdekoodin rajapinnat Yhdysvaltain ja EU:n talouden tietokantoihin. Euroopassa talouden tietokantoja tuottaa Eurostat ja Yhdysvalloissa The Bureau of Economic Analysis. (Ansip 2016, The Bureau of Economic Analysis 2016.)

Tilastollisen lähdekoodin ohjelmointikieli R tarjoaa integroidun pääsyn alueellisesti ja toimialoittain EU:n ja USA:n talouksien tietoihin kuten BKT:n, väestön, työllisyyden ja käytettävissä olevien tulojen tunnuslukuihin. Talouden tietokantojen avaaminen tuo todennäköisesti paljon synergioita yhteiskunnallisesti, liiketoiminnan kehittämisen, sekä tutkimuksen ja tuotekehityksen näkökulmasta. (Ansip 2016, The Bureau of Economic Analysis 2016.)

Euroopan sisämarkkina-alueella on Euroopan Komission toimesta käynnissä mittavia datainfrastruktuuriin sekä pilvipalveluun liittyviä hankkeita. Euroopassa tuotetaan tieteellistä dataa koko maailman mitataavassa eniten. Tieteellinen data on kuitenkin ollut tähän mennessä hajallaan. Komission tavoitteena on kerätä tieteellinen data yhteen paikkaan pilvipalveluun, ensin tieteellisen yhteisön käytettäväksi, mutta tulevaisuudessa myös julkisyhteisöjen ja yritysten vapaasti käytettäväksi. Hanke voi luoda kokonaan uusia teollisuuden aloja. Avoimen tutkimusdatan käytöstä voidaan saavuttaa Euroopassa valtavat taloudelliset hyödyt. (Euroopan komissio 2017.)



### 3.4 Johtopäätökset avoimen datan taloudellisista vaikutuksista

Avoimen datan hyödyntämisessä voidaan saavuttaa taloudellisia etuja, jos sitä käyttävillä organisaatioilla on tarpeeksi osaamista, resursseja, pääomaa ja mahdollisesti myös hyvät verkostot avointa dataa tarjoavien organisaatioiden kanssa. Datan avaaminen saattaa tuoda hyvinkin paljon taloudellisia etuja myös yksityiselle sektorille, jos avoimen tiedon käyttöön panostetaan ja sitä osataan hyödyntää palveluiden, tuotteiden ja innovaatioiden kehittämisessä. Yksityinen sekä julkinen sektori tarvitsivatkin opastusta ja tukea avoimen datan käytössä, jota valtio voi omalla avoimen datan käyttöönottoon ja hyödyntämiseen panostavalla toiminnallaan edesauttaa. Lopulta avoimen datan käyttö voi edistää työllisyyttä, tuottavuutta ja kasvua ja näin ollen se voi olla koko kansantalouden kannalta tärkeä tekijä.

Kustannushyötyjen mittaamista ja niiden tulosten arviointia tulee kehittää ja seurata. Tulokset auttavat esimerkiksi valtiota ja yrityksiä arvioimaan kannattaako avoimen datan avaamista edistää ja mihin tuki kannattaa kohdistaa ja miten.

Julkishallinnon avoimen datan läpinäkyvyydellä voidaan saavuttaa kansantaloudellisia hyötyjä mm. toimintoja tehostamalla, verotuloja kasvattamalla ja päällekkäisyyksiä karsimalla. (VM15a/2015, 19) Julkishallinnon datan avaamiselle on asetettu tavoitteeksi vuosi 2020, mutta osa organisaatioista pitää aikataulua liian kireänä (VNK 2017, 34). Datan avaaminen on pitkäjänteinen prosessi, jota saattaisi edesauttaa välitavoitteiden asettaminen. Panostaminen kohteisiin, joista arvioidaan olevan yhteiskunnallisesti ja taloudellisesti eniten hyötyä, on järkevin tapa edetä datan avaamisen osalta.

Kansainväliset kehityshankkeet ja Euroopan Komission aktiivinen toiminta avoimen datan avaamisessa ja käytön edistämisessä, tuovat mitä todennäköisemmin mukanaan uusia innovaatioita ja taloudellisia hyötyjä, sekä laajentavat avoimen datan käyttöä. Suomenkin kannattaa olla aktiivisena toimijana ja osallistujana EU:n kehityshankkeissa.

## 4 KÄYTTÖÖNOTETTAVA AVOIN TALOUSARVIODATA

Opinnäytetyön käytännön osuudessa toteutetaan Valtion avoimen talousarviodatan käyttöönotto ja jatkojalostaminen. Tavoitteena on tehostaa ja kehittää ennustetoimiston julkisen talouden ennuste ja seuranta työtä. Tavoitteiden toteutuessa seurannaisvaikutuksena analyysien ja ennusteen tuotanto nopeutuu sekä aineiston laatu paranee. Tämä puolestaan voi tuoda mukanaan kustannustehokkuutta ja taloudellisia hyötyjä.

Ennen avoimen talousarviodatan käyttöönottoa kuvaillaan lyhyesti, mitä talousarvioalla tarkoitetaan ja miten aineistoa on käytetty tähän mennessä ennustetoimistossa. Tämän lisäksi käydään läpi käyttöönotettavan avoimen talousarviodatan ominaisuuksia, ja sitä mitä tietoa datasta halutaan ottaa käyttöön, sekä sitä miten se täyttää avoimelle datalle asetetut vaatimukset.

### 4.1 Valtion talousarvio

Valtion talousarvio on valtion taloutta ja varainhoitoa koskeva suunnitelma, jonka avulla arvioidaan valtion vuotuiset tulot ja menot. Tuloja tarkastellaan pääluokittain eli osastoittain, esimerkiksi verotulojen mukaan. Pääluokan tulot jakaantuvat hienojakoisemmalla tasolla tulolukuihin- ja momentteihin. Vastaavasti menot jakaantuvat omiin pääluokkiinsa, eli ministeriöiden eri hallinnonaloihin. Hallinnonalojen menot jakaantuvat yksityiskohtaisemmalle tasolle menolukuihin ja menomomentteihin. (vm.fi n.d..)

Talousarvion laadintaprosessi etenee niin että ensin Valtiovarainministeriö tekee ehdotuksen seuraavan vuoden budjetista. Tämän jälkeen ehdotus käsitellään hallituksen budjettiriihessä, jossa laaditaan talousarviosta hallituksen esitys. Hallituksen esitys käsitellään eduskunnassa, joka päättää seuraavan vuoden talousarviosta. Budjetista laaditaan eduskunnan kirjelmä. (vm.fi n.d..)

### 4.2 Talousarviodatan käytön lähtötilanne Suomen Pankissa

Tähän asti valtion talousarviodatasta on päivitetty ennustetoimistossa Excel-taulukkoa karkealla tasolla, pääluokittain tulojen ja menojen osalta. Käsien kerätystä datasta ei ole saatu kaikkea mahdollista tietoa ennustamisen- ja seurannan tueksi. Datasta on julkaistu visuaalisia näyttöjä Valtiovarainministeriön tutkibudjettia.fi-internetsivustolla. Sivuston visualisoinnit on esitetty aina viimeisimmästä talousarviosta, eikä vertailua esimerkiksi edellisiin vuosiin ole tehty. Sivulla ei myöskään ole visuaalisessa muodossa mm. lisätalousarvioita, hallituksen esityksiä sekä Valtiovarainministeriön ehdotuksia. Sivusto on myös

melko uusi, joten sitä ei ole tämänkään vuoksi hyödynnetty. Internetistä löytyvä visuaalinen informaatio valtion budjetista ei tällä hetkellä mm. edellä mainittujen syiden vuoksi täytä täysin datan syvällisemmän analyysin asettamia vaatimuksia.

Valtiovarainministeriön tarjoama Valtion avoin talousarviodata sen sijaan kattaa paremmin tarpeet seuranta- ja ennustetyön näkökulmasta. Lisäksi koneluettavassa muodossa oleva data on helpommin siirrettävissä hyödynnettäväksi eri järjestelmiin ja ohjelmiin. Näin ollen data on sopivaa aineistoa siirrettäväksi toimiston käyttöön tutkittavaksi ja testattavaksi myös Power BI-ohjelmalla.

#### 4.3 Tietotarpeiden määrittely

Avoimen talousarviodatan käyttöönoton alkuvaiheessa tuli päättää, mitä tietoa budjetti.vm.fi-sivustolta kerätään Suomen Pankin ennustetoimiston käyttöä varten. Tämä edellytti, että Suomen Pankin ennustetoimiston tietotarpeet valtion talousarviodatan osalta tuli määritellä. Tarvemääritykset tehtiin yhdessä julkiseen talouteen perehtyneen asiantuntijan kanssa. Ykkösprioriteetilla käyttöönoton osalta olivat Eduskunnan kirjelmät ajankohdalta 2015–2017. Eduskunnan kirjelmät, eli talousarviot ovat aina lopullisia päätöksiä tulevalle kaudelle budjetoiduista tuloista ja menoista. Käyttöön otettavat kirjelmät käsittävät valtion budjetit kolmen vuoden ajalta, mikä mahdollistaa vertailun eri vuosien välillä. Opinnäytetyössä käsitellään lähinnä tämän aineiston käyttöönottoa ja jatkojalostamista Power BI:lla.

Eduskunnan kirjelmien osalta päädytään siihen, että kaikki tarjolla oleva aineisto vuosilta 2015–2017 otetaan käyttöön. Näin ollen datan sisältömäärittelyyn ei käyttöönottovaiheessa ollut tarvetta tehdä rajauksia. Käyttöön otettavan datan määrittelyn jälkeen suunniteltiin, miten käyttöönotto toteutetaan Power BI raportointi- ja analyysivälineellä.

#### 4.4 Kuinka hyvin avoin talousarviodata täyttää avoimen datan vaatimukset?

Valtion talousarviodatan tiedot on julkaistu avoimen datan yleisessä datakatalogissa Avoindata.fi-palvelussa. Tiedot budjetti.fi.vm-sivustosta, jossa avoin data sijaitsee, löytyvät Valtiovarainministeriön internetsivuilta hyvin. Data on ladattavissa budjetti.fi.vm-sivustolta XML- ja CSV-muodossa 36 tunnin kuluttua siitä kun talousarvio on julkaistu. CSV-muodossa olevasta datasta on jokaisen vuoden aineistosta tehty omat CSV-tiedostot. Aineistoissa on mukana myös vertailuluvut edellisiltä vuosilta. CSV-tiedostot on nimetty järkevästi, niin että niitä hyödynnettäessä tiedostonimiä voidaan hyödyntää datan tuonnissa eri järjestelmiin. Data on lisensoitu kansainväliselle Creative Commons 4.0

(CC BY 4.0) -lisensille, jota tulee käyttää avoimen datan yhteydessä. (budjetti.vm.fi n.d., avoindata.fi 2015.)

Talousarviodata täyttää hyvin avoimen datan vaatimukset. Se on julkista, maksutonta, uudelleen käytettävää, sen tiedot on julkaistu avoimen datan katalogissa ja sen tiedot on helposti löydettävissä. Talousarviodatan käyttöehdot on määritelty avoimen datan CC BY-lisensillä. Datan hyödyntäjälle datan XML-formaatin tulkinta on haastavaa. Rakenteisen XML-formaatin käytöstä tulisikin antaa ohjeita. Tämä edistäisi datan jatkokäyttöä.

## 5 POWER BI

Ennustetoimiston projektissa valittiin talousarvio datan tiedonsiirtoa, käsittelyä ja jalostusta varten Power BI -niminen ohjelma. Se on Microsoftin tuottama Business Intelligence -väline, jota ei ole aikaisemmin käytetty Suomen Pankin ennustetoimistossa. Case-projektin avulla olikin oiva tilaisuus tutustua uuden välineen käyttöön ja sen tuomiin mahdollisuuksiin.

### 5.1 Business Intelligence ja Power BI

Business Intelligencelle (BI) ei ole yksikäsitteistä selitettä. Perusajatuksena BI:ssa on kuitenkin se, että tietoa kerätään mm. liiketoiminnan kehittämisen tueksi. Business Intelligence tukee päätöksentekoa ja näin myös kilpailua. Loppukäyttäjän välineinä tiedon esittelyssä ovat raportointi- ja analysointiratkaisut, mittaristot jotka esittelevät tunnuslukuja, web-portaalit sekä mobiilisovellukset. Menetelmien tavoitteena on se, että täsmällinen, oikea ja ymmärrettävä tieto saavuttaa oikean henkilön oikea-aikaisesti. (Hovi, Hervonen, & Koistinen 2009, 73.)

Business Intelligence -välineitä on markkinoilla paljon ja Power BI -ohjelma on yksi näistä tällä hetkellä tarjolla olevista tuotteista. Power BI on Gartnerin 2017 maaliskuussa julkistetun tutkimuksen mukaan kilpailukykyisimpiä BI-ohjelmia, mitä markkinoilla on tarjolla (Kuva 4) (Gartner 2017). Power BI:sta on ollut tarjolla vuodesta 2013 lähtien erilaisia versioita ja sen komponentteja on ollut yhdistettynä myös Exceliin. Power BI on suhteellisen tuore tuote Business Intelligence välineiden markkinoilla. Sillä on potentiaalia kehittyä hyvinkin täysve-riseksi Business Intelligence -ohjelmaksi, käynnissä olevan aktiivisen tuotekehittelyn avulla. Nykyisen kaltaisena versiona Power BI on ollut käytössä vuodesta 2015 lähtien.



Kuva 4. Microsoftin Power BI-markkina asema (Gartner 2017). (Täydennetty kuvakaappaus.)

## 5.2 Power BI pähkinänkuoressa

Power BI on raportointi- ja analytiikkaväline. Ohjelmalla esitetään dataa vuorovaikutteisesti mm. erilaisia mittaristoja käyttäen. Välineen käyttöönottoon on matala kynnyks myös loppukäyttäjille. Sen käytössä on monia tuttuja Excel-taulukkolaskentaohjelmaan liittyviä elementtejä, kuten Power BI:n Desktopin datan käsittely. Power BI on kuitenkin oma itsenäinen palvelu, joka ei edellytä Excelin asentamista. Power BI koostuu kolmesta osa-alueesta: Power BI -työpöytäsovelluksesta (Desktop), Power BI -pilvipalvelun julkaisualustasta (Power BI Service) ja Power BI Mobiilista. Power BI Mobiili -palvelua ei tässä työssä käsitellä. (hexcelligent.fi 2016.)

Microsoftin Power BI tuoteperhe on rakennettu Power BI Free-, Pro ja Premium tuotekokonaisuuksista, jotka tarjoavat käyttäjille erilaisia Power BI:n käyttöön liittyviä vaihtoehtoja. Free on ilmaisipaketti, jonka kuka tahansa voi ottaa käyttöön. Se tarjoaa mahdollisuudet henkilökohtaiseen käyttöön, mutta ei raporttien jakeluun, eikä ryhmäkäyttöön. Pro on kohdistettu pienemmille organisaatioille. Premium on tarkoitettu isommille organisaatioille, mutta se ei tarjoa jakelupalvelua. Käytännössä tämä tarkoittaa sitä, että isomman organisaation tulisi

hankkia sekä Pro, että Premium -versiot. Nämä kaksi tuotekokonaisuutta tarjoavat parhaat mahdolliset hyödyt ja ominaisuudet isoille organisaatioille. (hexcelligent.fi 2016.)

### 5.2.1 Power BI Desktop

Power BI Desktop on ohjelma, jossa tehdään datan pääasiallinen käsittely ja jalostaminen. Siihen voidaan tuoda dataa useasta eri lähteestä, mm. internetistä, avoimista tietolähteistä ja pilvipalveluista. Tietolähteistä tuotua dataa pystytään yhdistelemään, datalle voidaan tehdä datan analysointia, laskentaa, kyselyjä ja sillä voidaan luoda visualisoituja raportteja. Lisäksi työpöytään kuuluvat tietomallit, jotka ovat olennaisessa roolissa dataa yhdisteltäessä ja isoja tietojoukkoja käsiteltäessä. (hexcelligent.fi 2016; Microsoft 2017.)

Desktopin datan käsittelyssä on hyvin paljon samoja ominaisuuksia ja toimintatapoja kuin Excelinkin käytössä. Desktopiin on myös yhdistetty Excelin apuohjelmia, jotka ovat sulautuneet Power BI:hin sujuvasti käytettäväksi. Power BI:hin yhdistetyt Excelin apuvälineitä ovat Power Query, jolla muokataan ja muunnetaan dataa, Power Pivot, jolla yhdistellään tauluja ja tehdään niistä laskentaa, sekä Power View -kaaviot. (hexcelligent.fi 2016.)

### 5.2.2 Power BI Service ja Dashboard

Desktopilta valmiit raportit ja data siirretään julkaisemalla (Publishing) ne pbix-tiedostona Power BI -pilvipalveluun (Power BI Service). Eri käyttäjäryhmillä voi olla ryhmäkohtaisia käyttöoikeuksia pilvipalvelussa olevaan aineistoon.

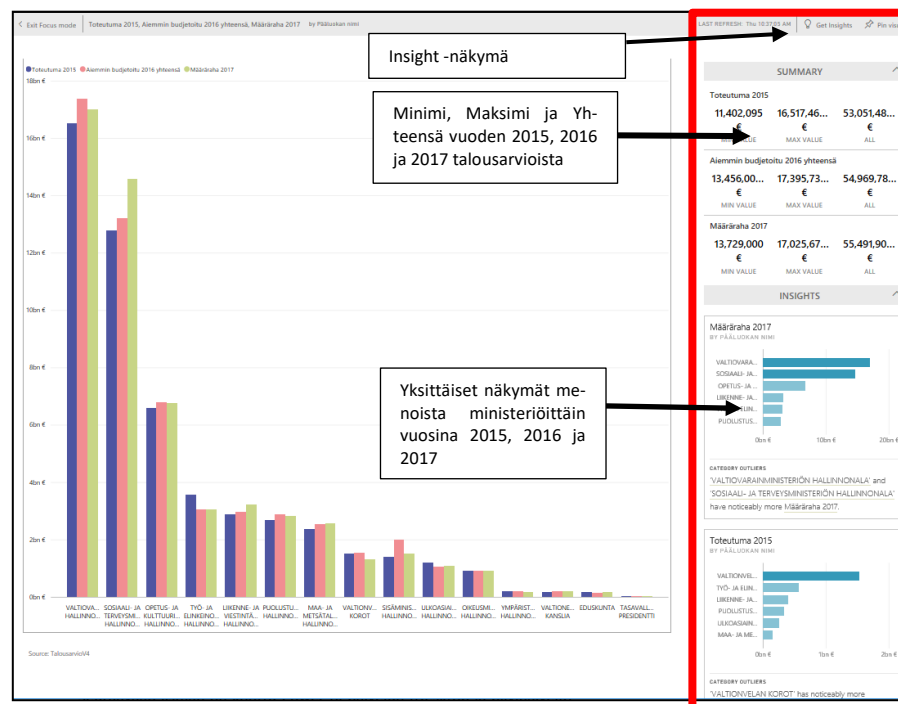
Raportit on mahdollista julkaista ja jakaa sähköpostilinkkinä, internetissä ja sosiaalisessa mediassa, esimerkiksi Twitterissä. Jaetut raportit ovat siis kenen tahansa käyttäjän käytettävissä, eikä niiden käyttö edellytä kirjautumista Office365-palveluun. Raportteja voi hyödyntää selaimella sekä mobiiliapplikaatioilla.

Power BI -pilvipalvelussa on mahdollisuus käyttää joko yrityksen omia (vaatii Pro-lisenssin), tai eri pilvipalveluiden tarjoamia valmiita sisältöpaketteja (Content Packs). Sisältöpakettien käyttäjillä tulee olla tunnukset siihen pilvipalveluun, jonka sisältöpaketteja käyttää. Paketteihin on luotu valmiita mittareita, koontinäyttöjä, tietomalleja ja raportteja, joiden avulla on Power BI:ssa mahdollista helposti analysoida eri lähteistä saatavaa dataa. Google Analyticsin sisältöpaketilla voidaan esimerkiksi kerätä valmiiden mittaristojen raporttien, sekä koontinäyttöjen avulla informaatiota verkkosivujen kävijäseurannasta. (hexcelligent.fi; Laksola & Mauno 2017, 27.)

Dashboard on Power BI:n pilvipalvelun julkaisualustalle rakennettava koontinäyttö, jolle on mahdollista kerätä kaavioita useasta eri raporttiedostosta. Koontinäytön tavoitteena on toimia nopean päätöksen välineenä. Sen tavoitteena on antaa yhdellä silmäyksellä kokonaisvaltainen näkökulma jostakin ilmiöstä tai asiasta. Koontinäytön jakaminen on tällä hetkellä mahdollista vain Office365-pilvipalveluun kirjautuneille Pro-lisenssiä käyttäville henkilöille. (Laksola ym. 2016, 25.)

Ennustamisen ja seurannan näkökulmasta tärkeimmistä seurattavista talousarvion indikaattoreista on mahdollista rakentaa koontinäyttö. Se antaa yleiskuvan esimerkiksi julkisen talouden tulo- ja menoarvion yleisestä tilanteesta. Koontinäytölle voidaan tuoda esimerkiksi vuoden 2017 tulo- ja menoarviota kuvattaessa visualisoinnit, jossa kuvataan yhdellä kaaviolla kokonaismenoja ja toisella kokonaistuloja. Mitä kaikkea muuta menoihin tai tuloihin liittyvää sitten halutaankin tarkastella tarkemmin, edellyttää se siirtymistä kaavion kautta raporttiedostoon, joka ei siis sijaitse koontinäytössä.

Koontinäyttö tarjoaa yksittäisestä kaaviosta myös Insights-näkymän, jossa sen sisältöä on mahdollista analysoida monipuolisemmin. Alla olevassa talousarvion menoja ministeriöittäin kuvaavassa kaaviossa esimerkiksi saadaan automaattisesti Insight -näkökulman kautta yhteenvetotietoa budjettimäärärahojen minimi-, maksimi -ja yhteenlasketuista arvoista. Tämän lisäksi näkökulma mahdollistaa tarkastelun budjettimäärärahoista eri vuositasoilla (Kuva 5). Kaiken tämän Power BI osaa tehdä automaattisesti, älykästä päättelyä käyttäen.



Kuva 5. Dashboardin Insight-näkökulma. (Täydennetty kuvakaappaus.)



## 6 TALOUSARVIODATAN KÄYTTÖNOTTO POWER BI:LLA

Ennen varsinaista datan käyttöönottoa ja työskentelyn aloittamista talousarviodatalla, käydään läpi miten aineisto tuodaan Power BI Desktopiin, mitä yleisiä työvaiheita kuuluu datan käyttöönottoon ja jatkojalostamiseen, sekä mitä on hyvä huomoida datan laadun ja käsittelyn osalta.

### 6.1 Työvaiheet Desktopilla

Aineistoa tuotaessa Power BI Desktopiin tietokannoista tai ulkopuolisista lähteistä, aineiston käsittelyyn ja jatkojalostamiseen kuuluu yleensä tehtäviä, jotka etenevät vaiheittain.

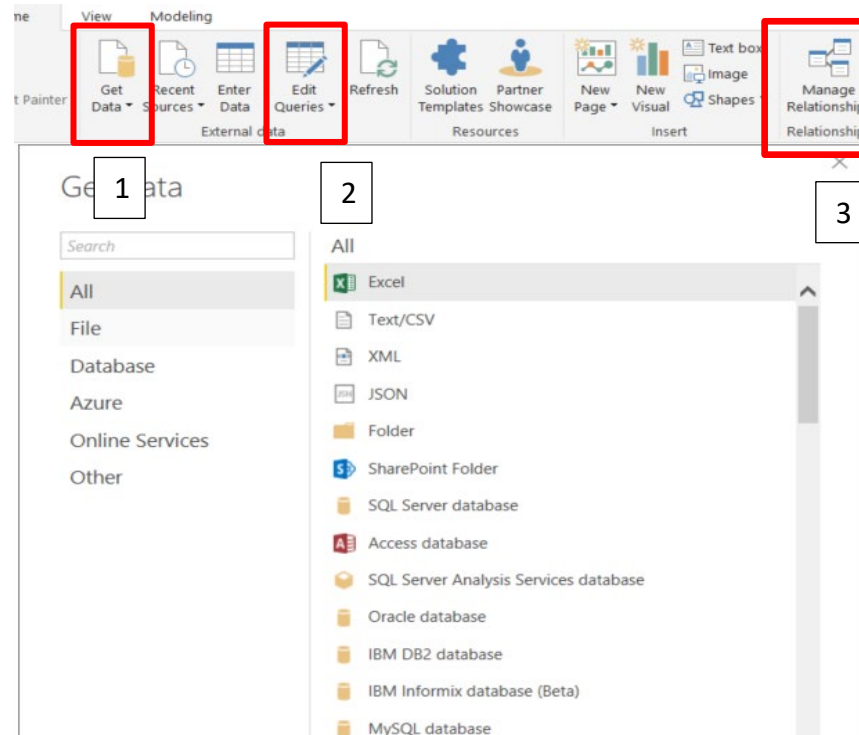
Ensimmäisenä työvaiheena on aineiston lataus, joka tehdään Get Data-toiminnolla. Aineisto ladataan omista tietokannoista, tai ulkoisista tietolähteistä (Kuva 6, kohta 1). Power BI tukee 70 eri tiedostomuotoa, joten tiedon tuomisessa Power BI:hin ei yleensä esiinny ongelmia.

Ensimmäisen vaiheen jälkeen tulee data jalostaa sellaiseen muotoon että se vastaa käyttäjien tarpeita. Aineiston muokkaaminen ja kyselyiden tekeminen tehdään kyselyeditorissa (Edit Query) (Kuva 6, kohta 2). Myös Data-näkymässä (kuva 7, kohta B) on mahdollista tehdä vastavia toimintoja, kuten laskentaa, mutta muokkausvaihtoehtoja on tässä näkymässä vähemmän.

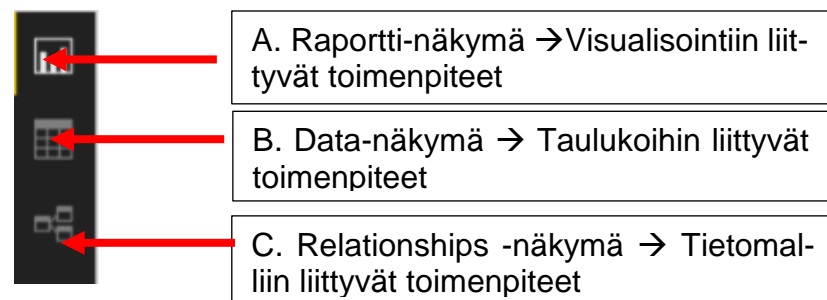
Yhdessä pbix-tiedostossa käsitellään yleensä useita tauluja, joten datan siivoamisen jälkeen tulee selvittää taulujen väliset keskinäiset yhteydet. Manage Relationships -näkymän tietokantataulujen katseluun ja muokkaamiseen siirrytään Desktopin valintanauhasta (Kuva 6, kohta 3), tai Desktopin vasemmassa reunassa olevan Relationships-näkymän kautta (Kuva 7, kohta C).

Datan taulujen yhteyksien selvittämisen jälkeen aineistosta on tapana tehdä analyysiä tai esimerkiksi esitysaineistoa varten visualisointeja. Visualisoinnit luodaan Raportti-näkymän (Kuva 7, kohta A) raporttisivuille.

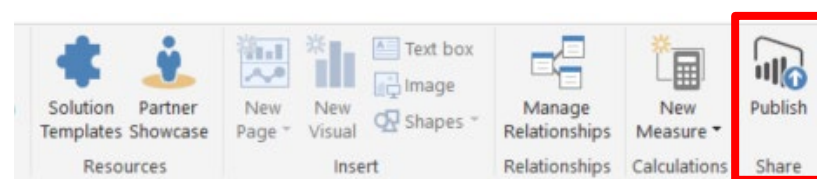
Visualisoidut raportit ja data-aineistot halutaan usein jakaa esimerkiksi organisaation sisällä tai ulkopuolisille tahoille eteenpäin. Tällöin raporttiedosto tulee julkaista Power BI:n pilvipalvelussa Power BI Servissä. Tiedoston julkaiseminen Power BI:n pilvipalvelussa tehdään Desktopin valintanauhalla Publishing-valinnalla (Kuva 8). (hexcelligent.fi, 2016; Microsoft 2017.)



Kuva 6. PowerBI:n työnkulut. (Täydennetty kuvakaappaus.)



Kuva 7. Datan katselunäkymät. (Täydennetty kuvakaappaus.)



Kuva 8. Tiedoston julkaiseminen Power BI -pilvipalvelussa. (Täydennetty kuvakaappaus 3.)

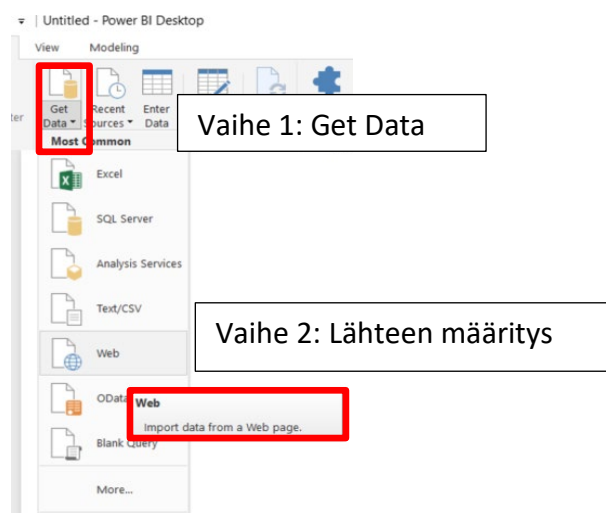
## 6.2 Datan käyttöönoton toteutus Power BI:lla

Aineiston tuonti Power BI Desktopiin on periaatteessa hyvin helppoa jos datalle ei tarvitse tehdä mitään käsittelyä. Power BI:n voi yhdistää monipuolisesti eri tietolähteisiin. Tietolähteet voidaan jakaa kolmeen eri ryhmään: tiedostoihin, sisältöpaketteihin ja tietokantoihin. Tiedostomuotoja voivat olla mm. tekstityyppiset CSV- tai XML-muotoiset tiedostot. (Microsoft 2017; Laksola ym. 2016, 27.)

Talousarviodatan osalta oli aluksi tarkoituksena luoda yhteydet XML-tiedostoihin, mutta käyttöönoton alkuvaiheessa ilmeni että yhteydenotto ei onnistunut Get Data -toiminnolla. Datan siirtoon liittyvät ongelmat saattoivat johtua VM:n XML-tiedostojen monimutkaisesta rakenteesta.

Tämä vuoksi päädyttiin Valtion talousarviodatan osalta CSV-muotoisen datan käyttöönottoon. Aluksi luotiin dynaamiset linkit talousarviodatan internetsivuille CSV-tiedostoihin, koska datan haluttiin päivittyvän mahdollisten muutosten varalta. Lisäksi linkkien avulla pystyttiin hyödyntämään datan kopiointia toiselle nimelle uusien aineistojen ilmesyessä avoimen datan budjetti.fi-internetsivuille.

CSV-muotoinen data noudettiin sivuston URL-osoitteen avulla, Power BI:n Get Data -toiminnolla (Kuva 9, vaihe 1), jonka jälkeen vielä määriteltiin lähteen tyyppi, joka tässä tapauksessa oli Internetin-lähde (Web) (Kuva 9, vaihe 2).



Kuva 9. Datan tuonti Power BI:n Desktopiin Get Data -toiminnolla. (Täydennetty kuvakaappaus.)

### 6.3 Datan sisällön tarkastus

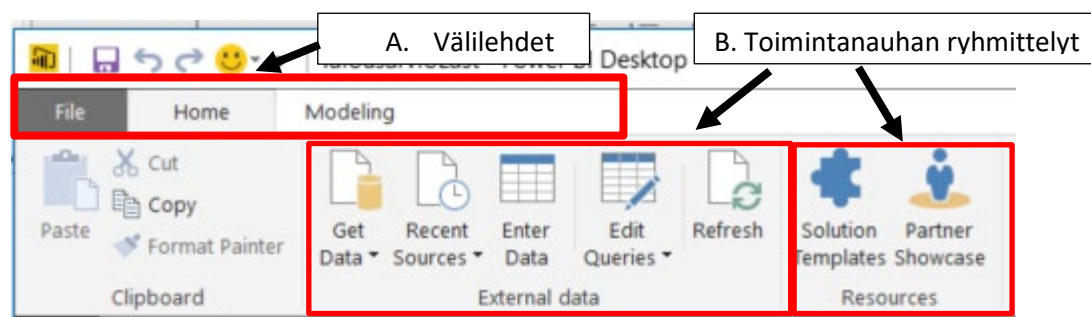
Datan siirron jälkeen tarkastettiin aineisto ja käsiteltiin se niin, että data oli käyttökelpoista ja vastasi käyttötarpeita. Datan laadun varmistuksessa talousarvioidatan osalta tarkastettiin mm. tietomallissa olevien taulujen väliset yhteydet, kenttien tietotyyppien oikeellisuus, missä muodossa tiedot esitettiin, ylimääräiset sarakkeet, jotka haluttiin poistaa tai vastaavasti sarakkeet, jotka haluttiin yhdistää, sekä kentissä oleva mahdollinen virheellinen data. Ulkopuolisesta lähteestä tuotavat taulukot saattavat sisältää myös tyhjiä rivejä tai sarakeotsakkeet saattavat olla riveillä, eivätkä otsikkotietona. Lisäksi oli huomioitava esimerkiksi, minkä kansallisuusasetuksen mukaan desimaalierottimet ja päivämäärämerkinnät oli määritelty.

Talousarvioidatan datan laatu oli hyvää ja sisältö vastasi melko pitkälle käyttötarpeita. Näin ollen edellä mainittuja muokkaustoimintoja ei lopulta tarvinnut juurikaan tehdä datan tarkastamisen jälkeen. Muutoksia tehtiin lähinnä tietotyyppien muunnosten, null-arvojen korvaamisen, sekä sarakkeiden lisäysten osalta.

## 7 TALOUSARVIODATAN KÄSITTELY POWER BI:LLA

Ennen talousarvioidatan käsittelyä Power BI:lla on hyvä kerrata, mitä dataan kohdistuvia käsittelyjä on mahdollista tehdä missäkin Power BI:n näkymässä. Report-näkymässä tehdään visualisointiin liittyvät toimenpiteet, Data-näkymässä käsitellään tietoa sisältäviä tauluja ja Relationships-näkymässä tietomalliin kohdistuvia tehtäviä. Kuvakaappaus näkymistä löytyy sivulta 19 (Kuva 7).

Jokaisessa näkymässä on omat välilehdet datan käsittelyyn liittyviin osa-alueisiin. Kuvassa 10 (Kohta A) on Data-näkymän välilehdet. Välilehdillä on valintanauha, joka mahdollistaa välilehdestä riippuen tauluun liittyvien toimintojen suorittamisen. Toimintanauhalla toimintakokonaisuudet on kerätty omiin ryhmiinsä (Kuva 10, kohta B).

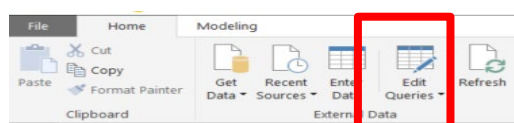


Kuva 10. Data -näkömän välilehdet ja ryhmittelyt. (Täydennetty kuvakaappaus.)

### 7.1 Talousarvion taulujen käsittely Power BI:lla

Talousarvion datataulujen muokkaustoiminnot, kuten tietotyyppien muutokset, laskenta, taulujen ja sarakkeiden yhdistelyt tehdään pääasiassa Power BI:n Edit Query -näkömässä. Edit Query -näkömään siirytään valintanauhan Edit Query -painikkeella (Kuva 11).

Kyselyeditorissa on paljon muunnos- ja muokkausmahdollisuuksia kuten tietotyyppien muunnosmahdollisuudet, sekä erottimien ja tekstin korvausmahdollisuudet jne. Lisäksi Edit Querylla voi poistaa sarakkeita ja rivejä. Sillä voi yhdistellä tauluja ja taulujen kenttiä tai tehdä laskentaa. Myös Data-näkömässä on mahdollista tehdä vastaavia toimintoja, kuten laskentaa, mutta muokausvaihtoehtoja on tässä näkömässä vähemmän.



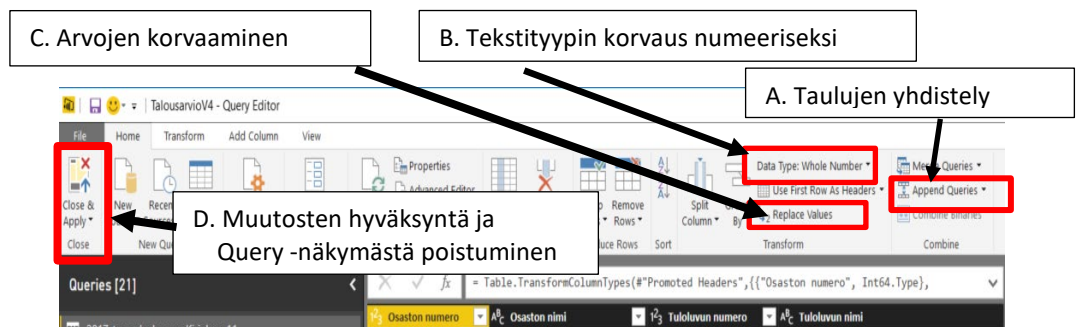
Kuva 11. Edit Query -näkömään siirtyminen. (Täydennetty kuvakaappaus.)

Tietomallissa jossa tarkastellaan mm. datataulujen välisiä yhteyksiä, Valtion talousarviodata sisälsi menojen osalta taulukot kaikista ministeriöistä. Tämä tarkoitti esimerkiksi menojen osalta sitä, että Power BI:hin olisi pitänyt määritellä yhteydet jokaisen viiteentoista ministeriön menotauluun erikseen. Menot -taulujen rakenteet olivat kuitenkin rakenteeltaan samat, kuten myös Tulot -taulujen rakenteet. Tämä helpotti loppuratkaisun tekemistä tietomallin taulujen yhdistämisessä Append Query -toiminnolla.

Vuoden 2017 talousarvion neljä tulotaulua yhdistettiin yhdeksi tulotauluksi, ja menotaulut yhdeksi menotauluksi. Append Queryn valinta tehtiin Edit Queryn valintanauhalla (Kuvio 12, Kohta A). Kun yhdistelmätaulukon lähdetaulukot päivittyvät, päivittyy myös yhdistelmätaulujen sisältö.

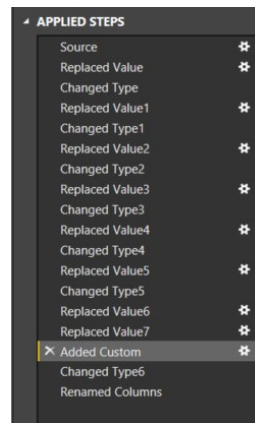
Lisäksi datan laskennassa käytettävien tietotyyppien sisältöä muokattiin niin, että tekstityyppiset kentät muutettiin kokonaisluvuiksi toiminnolla Data Type Whole: Number (Kuva 12, kohta B) ja Null-arvot korvattiin nolla-arvoiksi toiminnolla Replace Value (Kuva 12, kohta C).

Muokkaukset hyväksytään Close & Apply -valinnalla (Kuvio 12, kohta D). Toimenpiteen jälkeen Query Editori siirtyy takaisin taulukkonäkymään.



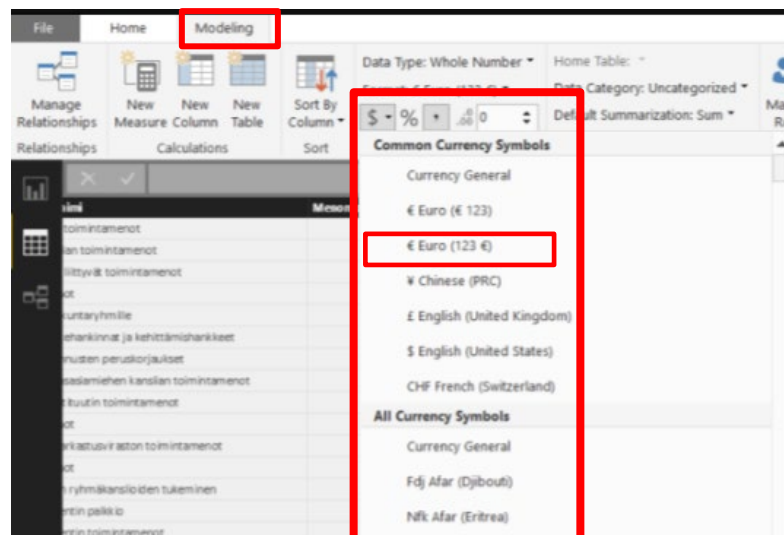
Kuva 12. Query Editorin valintanauha (Ribbon). (Täydennetty kuvakaappaus.)

Datan muokausvaiheessa työskentelyä helpotti se että kaikki Query Editorissa tehdyt muutokset näkyivät vaihe vaiheelta näkymän oikeassa reunassa olevassa Applied Steps -ikkunassa (Kuva 13). Muutokseen oli aina mahdollista palata ja tarvittaessa sen pystyi poistamaan. Poistojen vaikutukset tuli kuitenkin aina huomioida listalla oleviin seuraaviin muutoksiin.



Kuva 13. Applied Steps -ikkuna. (Kuvakaappaus.)

Valtion talousarvioidatan osalta oli vielä tarvetta muuttaa kokonaisluvuiksi muunnetut sarakkeet valuuttamääräisiksi sarakkeiksi. Valuuttamuunnokset tehtiin Data-näkymän Modeling-välilehdellä valintanauhasta \$-merkin alta pudotuslistalta (Kuva 14).

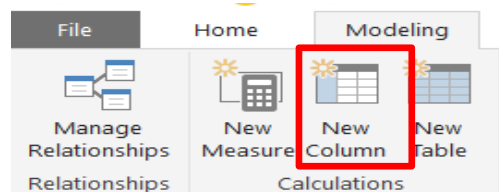


Kuva 14. Valuuttamuotoilu Taulukko-näkymän valintanauhassa. (Täydennetty kuvakaappaus.)

## 7.2 Laskenta Power BI:lla

Power BI mahdollistaa laskennan tekemisen Power BI:n Taulukko-, Edit Query- ja Visualisointi-näkymissä. Power BI:n Desktopissa käytetään laskennassa Data Analysis Expressions (DAX)-kaavakieltä. DAX käyttää hyvin samanlaisia toimintoja, operaatioita ja syntakseja, kuin Excelin kaavat. Laskennallisissa toimenpiteissä (Measures) hyödynnetään taulujen kenttälistoja, jotka lisätään kaavaan. Laskentakaavat nimetään omalla nimellään ja niitä voidaan tämän jälkeen hyödyntää myös muissa DAX-laskelmissa.

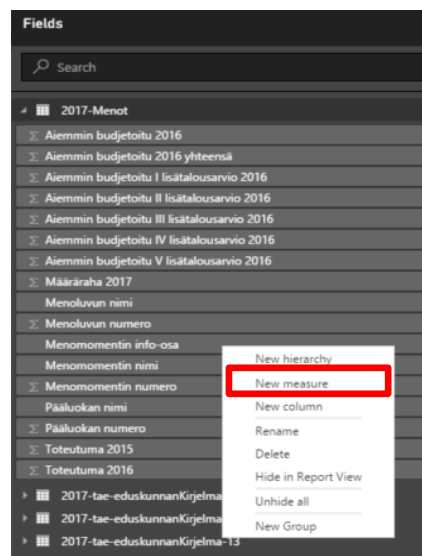
Talousarviodatan menot yhdistelmätaulussa laskettiin kaikki vuoden 2016 menot yhteensä, DAX-kaavaa käyttäen. Menot-tauluun luotiin uusi laskennallinen sarake (Custom column), johon yhteenlaskukaava tehtiin DAX-funktiolla. Laskentaa varten valittiin Data-näkymässä Modeling-välilehti ja sieltä Calculations-osiosta New Column -valinta, uutta laskennallista saraketta varten (Kuva 15). Seuraavaksi lisättiin Formula-kenttään DAX-kaava yhteenlaskutoimitusta varten.



Kuva 15. Laskennallisen sarakkeen lisääminen Modeling -välilehdellä. (Täydennetty kuvakaappaus.)

Tämän jälkeen DAX-funktiota, jossa oli laskettu kaikki 2016 vuoden menot yhteen, on mahdollista käyttää missä tahansa muussa laskutoimituksessa funktion nimellä [Menot 2016 Yhteensä].

Laskentaa on mahdollista suorittaa myös niin, että laskennalliselle kaavalle ei tehdä omaa laskennallista saraketta. Laskennallisesta kaavasta tehdään tällöin oma kenttä Fields-paneelissa, valitsemalla hiiren kaksoispainikkeen takaa valinta "New measure"(kuva 16). Tämän jälkeen kaava syötetään kaavaeditoriin, jonka jälkeen laskentakaava näkyy Fields-näkymässä omana kuvakkeenaan (Kuva 17).



Kuva 16. Fields-ikkunan New measure -valinta. (Täydennetty kuvakaappaus.)



Kuva 17. Kuva laskennallisesta kuvakkeesta. (Kuvakaappaus.)



## 8 TALOUSARVION PERUSRAPORTTIEN JA KOONTINÄYTÖN TOTEUTUS

Opinnäytetyön tavoitteena oli toteuttaa lopputuotteena perusnäytöt talousarvion raporteista sekä koontinäyttö. Tässä luvussa käsitellään raporttinäyttöjen ja koontinäytön toteutusta. Lisäksi käydään läpi raporttinäyttöjen julkaisemisen Power BI-pilvipalvelussa. Luvussa 9 arvioidaan lopputuloksia, sekä pohditaan paransivatko ne datan käyttöä ja visualisointien hyödyntämistä talousarvioidatan käytön nykytilanteeseen verrattuna.

Alussa on hyvä tietää raportti- ja koontinäyttöjen erot, koska molemmat näytöt voivat olla hyvin samankaltaisia näkymiä ulkoasullisesti. Perusajatus on se, että ensin luodaan Desktop-ohjelmalla kaaviot ja taulukot raporttinäyttöille. Yhteiskäyttöä ja koontinäytön rakentamista varten raportit julkaistaan pbix-tiedostossa Power BI Service -pilvipalvelussa. Koontinäyttöön on mahdollista kerätä useasta eri raporttietiedostosta (pbix-tiedosto) kaavioita. Koontinäytöltä kaaviota klikkaamalla päästään pilvipalvelussa olevaan raporttietiedostoon, josta löytyy enemmän tietoa kaavion sisällöstä, kuten kaavion datasta. Raporttinäyttöjä voi jakaa Internetissä, mutta koontinäyttöjen jakaminen on mahdollista vain ryhmätyötilojen ja käyttäjäryhmien välillä.

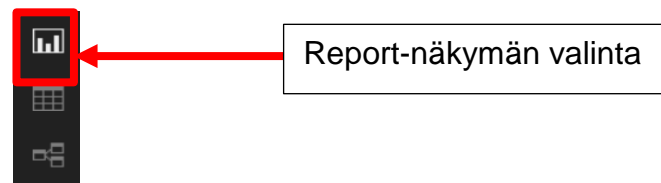
Talousarvioaineistosta oli tavoitteena rakentaa talousarvion tuloista ja menoista perusraporttinäytöt. Tuloja ja menoja halutaan seurata hienojakoisemmalla tasolla: menojen osalta pääluokittain menoluvun ja menomomentin mukaan. Tuloja tarkastellaan osaston nimen ja sen tulolukujen, sekä momenttien mukaan. Lisäksi seuranta halutaan tehdä eri budjettivuosien välillä, euromääräisinä tuloina ja menoina, sekä prosenttimuutoksien osalta. Talousarvion esittämisen osalta kohde-ryhmänä ovat asiantuntijat, jotka käyttävät raportteja työnsä tukena seuranta- ja ennustetyössä.

### 8.1 Raporttinäyttöjen visualisointien toteutus Power BI:lla

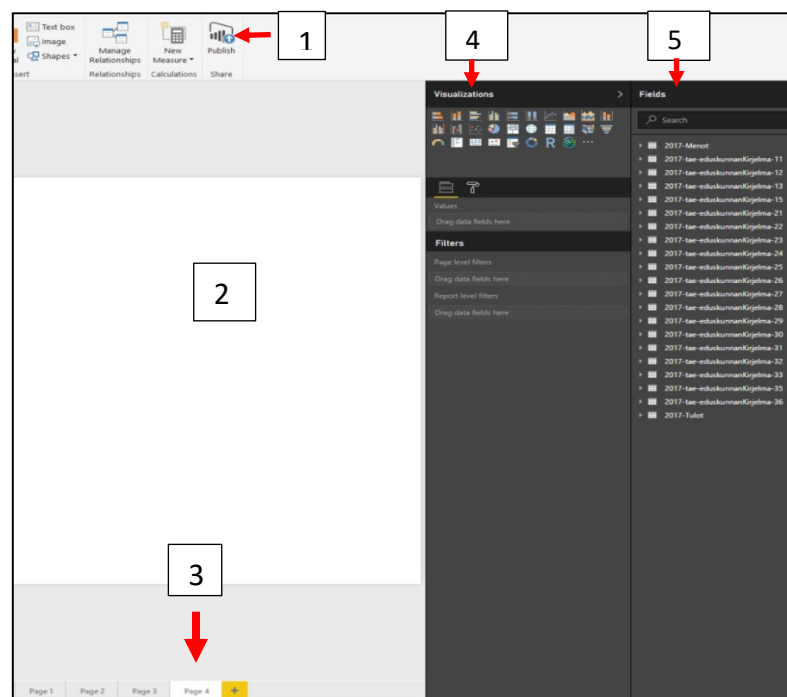
Talousarvion visualisoinnissa hyödynnettiin sekä perinteisiä, että uusia tarjolla olevia mahdollisuuksia. Raporttinäytöt koostuvat kaavioista sekä taulukoista. Perusraporttinäytöt rakennettiin pylväs-, vesiputous-, kortti- ja puukaavioita käyttäen. Lisäksi testattiin Microsoftin yhteistyökumppanien tarjoamia kustomoituja visualisointeja, sekä infografiikkaa. Toteutuksissa käytettiin myös osittajia (Slicer) ja hierakia näkymää.

Visualisointien toteutus on hyvin yksinkertaista ja nopeaa. Toteutus tapahtuu Report-näkymän valinnalla (Kuva 18). Näkymä jakaantuu viiteen työskentelyyn liittyvään perusosa-alueeseen: valintanauhaan, alustaan, visualisointi- ja kenttäpaneeliin (Kuva 21). Valintanauhalla valitaan raportteihin ja visualisointiin liittyvät tehtävät (Kuva 21, kohta 1).

Alustalla luodaan visualisointeja sekä järjestellään ja lisätään Sivuja (Page) -välilehti (Kuva 19, kohta 2 ja 3). Visualisointipaneelissa valitaan kaavioita, määritellään värejä ja X- ja Y-akselit, sekä lisätään suodatuksia jne. (Kuva 19, kohta 4). Kenttä -paneelissa valitaan taulujen kenttiä visualisointipaneeliin, esimerkiksi X- ja Y-akselien arvoiksi tai suodatimiksi (Kuva 19, kohta 5).



Kuva 18. Report -näkymän valinta. (Täydennetty kuvakaappaus.)

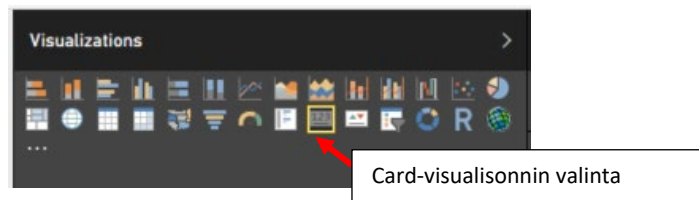


Kuva 19. Visualisointi- ja kenttäpaneelit. (Täydennetty kuvakaappaus.)

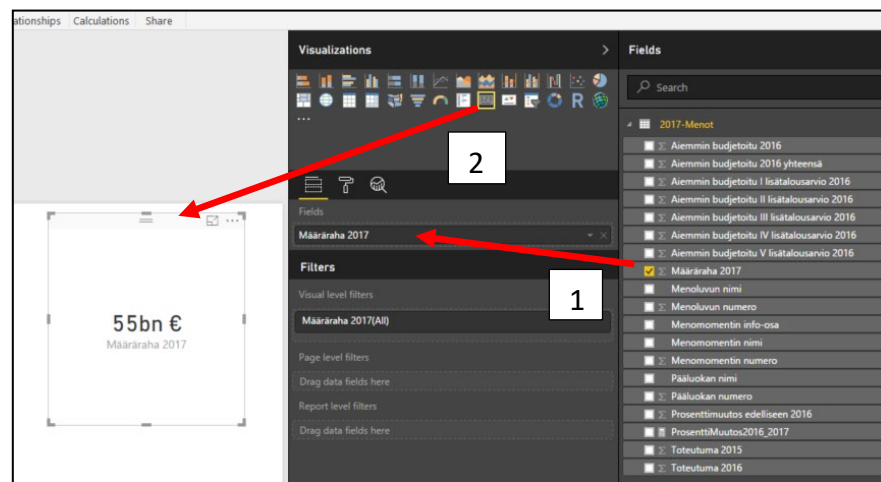
Talousarviodatasta luotiin ensin perusraporttinäkymät talousarvion tuloista ja menoista vuodelta 2017. Seuraavissa luvuissa kuvataan vain menomäärärahoihin liittyvät kaavioiden luomistehtävät. Tulojen osalta toimenpiteet ovat hyvin samankaltaiset, poikkeavuus on lähinnä kenttätiedoissa.

Menoista ensimmäisenä esitetään arvio käytettävissä olevista menomäärärahoista yhteensä. Luku on siis yksinkertainen numeroarvo, joka on mahdollista esittää esimerkiksi Kortti-näkymän (Card) avulla (Kuva 20). Aluksi valittiin Visualisointi-paneelista Card-kuvake, joka vedettiin Fields-paneelista raportti-sivulle. Tämän jälkeen valittiin Menot 2017 -taulusta kenttä "Määräraha 2017", joka siirrettiin Visualisointi-paneelin Fields-kenttään (Kuva 21, kohta 1). Kortti-visualisoinnilla syntyy yksinkertainen näkymä vuoden 2017 budjetoiduista kokonaismenoista

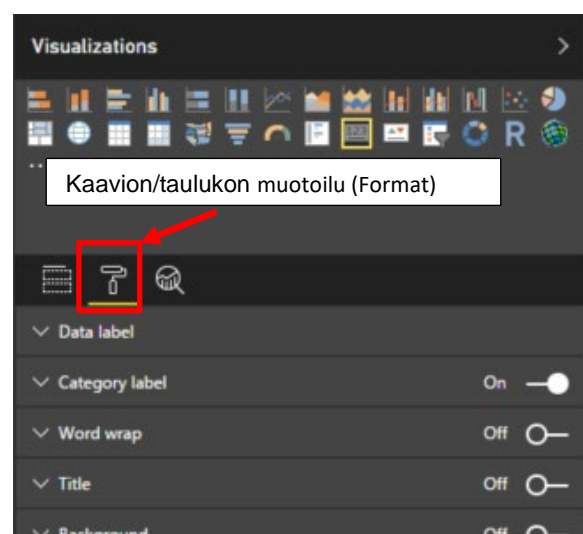
(Kuva 21, kohta 2). Valkoisella taustalla oleva kortti ei ulkoasullisesti näytä kovin näyttävältä, joten kortin ulkoasua muotoiltiin Visualisointi-paneelin Format-välilehdellä (Kuva 2). Kuvassa 23, näkyy lopputulos muotoillusta Kortti-kaaviosta.



Kuva 20. Card-visualisoinnin valinta Visualization-näkymässä. (täydennetty kuvakaappaus.)



Kuva 21. Määrärahat 2017 valinta Fields-paneelista Visualization-näkymän kentäksi sekä valmis Kortti-visualisointi. (Täydennetty kuvakaappaus.)



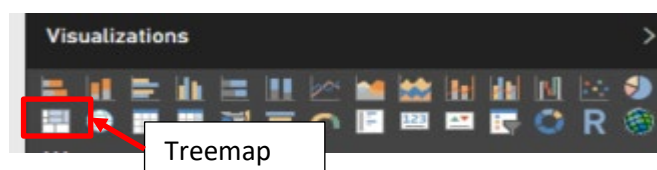
Kuva 22. Kaavion ja taulukon muotoilu Format-paneelissa. (Täydennetty kuvakaappaus.)



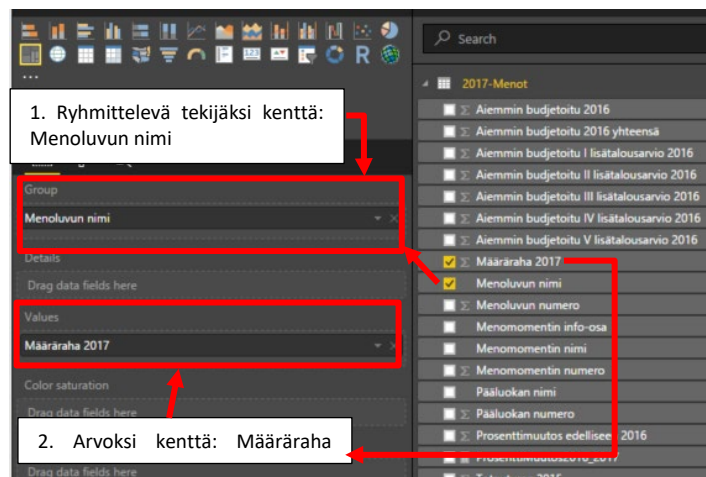
Kuva 23. Kortti (Card) -kaavio kokonaismenoista. (Kuvakaappaus.)

Talousarvion seurannassa oli tärkeää seurata menoja myös pääluokittain, menolukujen sekä sen menomomenttien mukaan. Menojen hienojakoisempaan seurantaan lukujen ja momenttien mukaan, sopii hyvin ruutupuu-kaavio eli Treemap (Kuva 27). Ruutupuu jakaa kokonaisuuden samanmuotoisina ruutuina hierarkiin osiin. Jokaisen ruudun koko kuvaa sen osuutta kokonaisuudesta. (Koponen, Hildén & Vapaasalo 2016, 237.) Menojen jakaantumista pääluokittain seurataan Slicerin eli osittajan avulla (Kuva 29). Pääluokkiin aineistossa kuuluvat eri hallinnonalat, kuten ministeriöt, eduskunta ja valtioneuvoston kanslia.

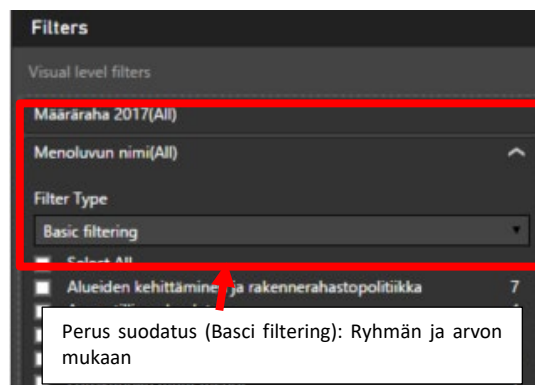
Ensimmäisenä toteutettiin ruutupuu siitä, miten määrärahat olivat jakaantuneet menoluvuittain. Ruutupuu valitaan Visualization-näkymästä (Kuva 24). Ruutupuun ryhmitteleväksi tekijäksi valittiin kenttä "Menoluvun nimi" (Kuva 25, kohta 1) ja arvoksi kenttä "Määräraha 2017" (Kuva 25, kohta 2). Power BI sijoittaa Visualizations-näkymässä automaattisesti oletusarvoisesti kenttiä ryhmään (Group) ja arvoihin (Values) (Kuva 25). Kenttiä on tarvittaessa myös mahdollista vaihtaa. Arvoja voidaan rajoittaa esimerkiksi ryhmän tai arvon mukaan perusuodatuksella (Filtering) (Kuva 26). Suodatuksessa on mahdollisuus käyttää myös edistynyttä suodatusta (Advanced filtering). Kuvassa 27, on nähtävissä lopputulos budjetoiduista määrärahoista menoluvuittain ruutupuu-visualisointina.



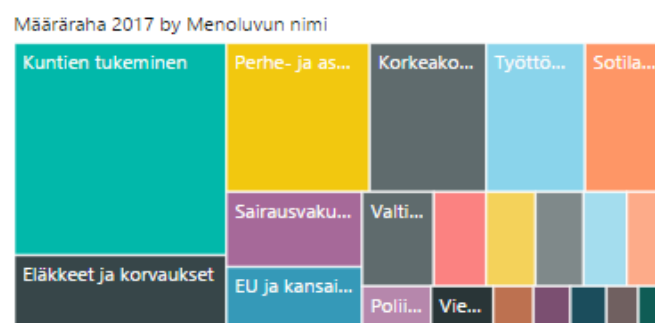
Kuva 24. Ruutupuu-visualisoinnin valinta. (Täydennetty kuvakaappaus.)



Kuva 25. Budjetoitujen menojen tarkastelu menoluvuittain. (Täydennetty kuvakaappaus.)

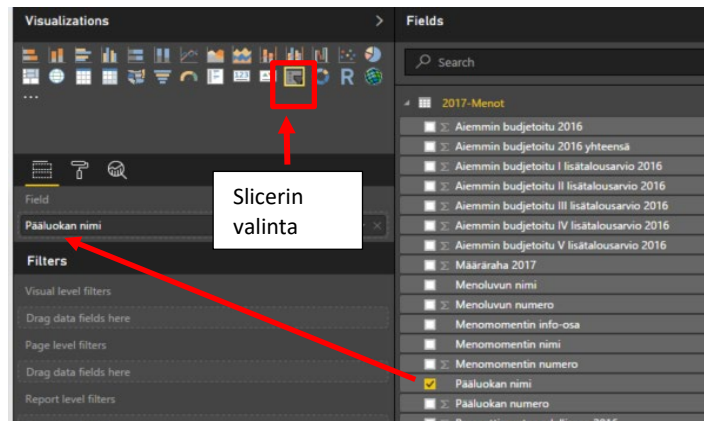


Kuva 26. Suodatusmahdollisuus perus suodatuksella (Basic filtering). (Täydennetty kuvakaappaus.)

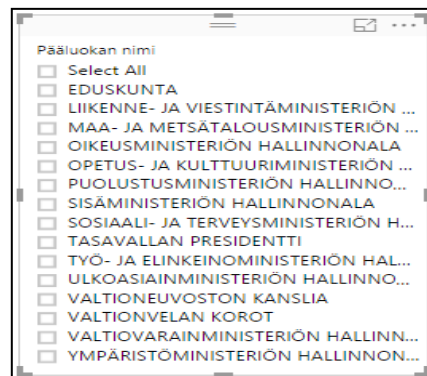


Kuva 27. Ruutupuu menoluvuittain. (Kuvakaappaus.)

Osittajan määrittelyn teko tapahtui seuraavasti. Valitaan Visualization-näkymästä Slicer ja osittajan kentäksi "Pääloukan nimi" (Kuva 28). Kuvassa 29 on nähtävissä lopputuloksena syntynyt Osittaja-näkymä kaikista pääloukista. Tämän jälkeen osittajan avulla on mahdollista analysoida, miten ruutupuu muuttuu menoluokittain, jos tarkastellaan esimerkiksi yksittäisen pääloukan, kuten Valtiovarainministeriön menoja.



Kuva 28. Osittajan (Slicer) valinta. (Täydennetty kuvakaappaus.)



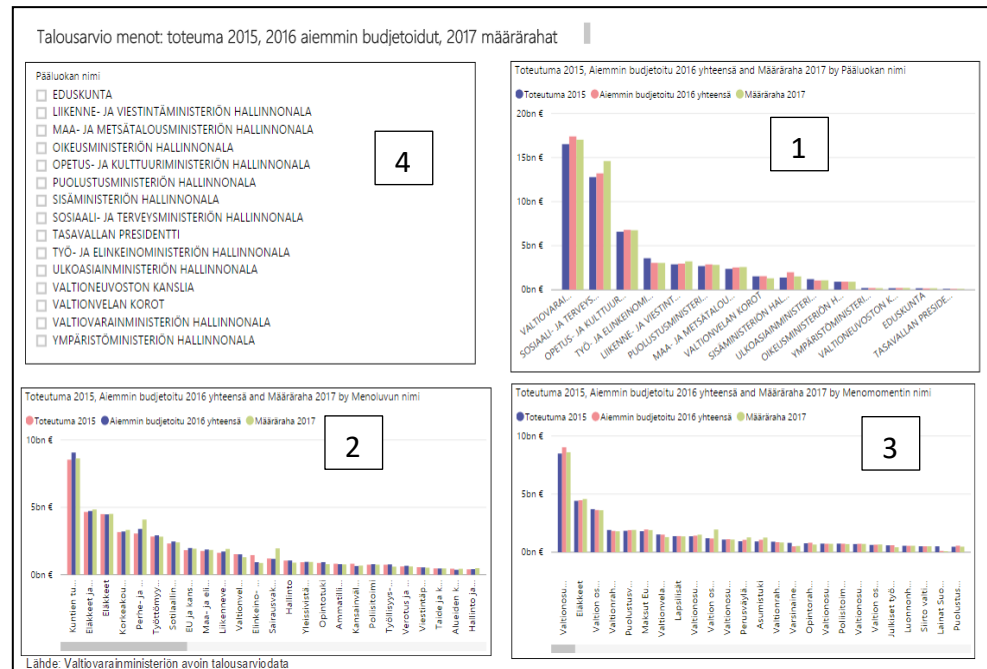
Kuva 29. Osittaja-näkymä Pääloukittain. (Kuvakaappaus.)

Lopuksi samaan raportti näyttöön yhdistetään luodut Card-visualisoinnit kokonaismenoista ja tuloista, ruutupuu-näyttö menoluvuista ja tulomomenteista, sekä osittajat osaston nimen ja pääloukan mukaan. Yhdistelmä-näyttö helpottaa menojen ja tulojen tarkastelun eri kriteerein samalla näytöllä (kuva 30).



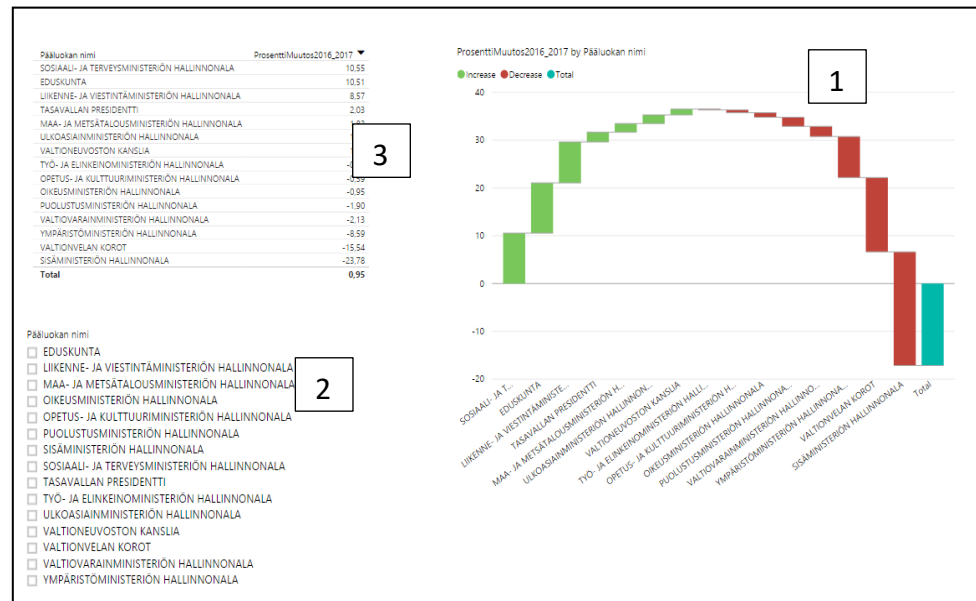
Kuva 30. Raportti-näkymä tuloista ja menoista joka mahdollistaa eri kriteerien valinnat. (Kuvakaappaus.)

Talousarvion seurannan tueksi toteutettiin myös yhdelle raportti-näytölle pylväskaaviot sekä osittaja, tulojen ja menojen eri vuositasokehityksen vertailua varten. Pylväskaaviot toteutettiin pääluokittain (Kuva 31, kohta 1), menoluvuittain (Kuva 31, kohta 2), sekä menomomenttien mukaan (Kuva 31, kohta 3), aikaväliltä 2015–2017. Osittajana käytettiin "Pääluokan nimi" kenttää (Kuva 31, kohta 4)).



Kuva 31. Osittaja pääluokista sekä pylväskaaviot menoista: pääluokan, menoluvun ja menomomentin mukaan. (Kuvakaappaus.)

Lisäksi seurannassa oli tarvetta tehdä eri vuositasojen prosenttimuutosten vertailua. Tätä tarvetta varten luotiin visualisointi vesiputouskaaviolla (Waterfall). Kaavion avulla vertaillaan edellisen vuoden menojen prosenttimuutoksia aikavälillä 2016–2017 (Kuva 32, kohta 1). Yksityiskohtaisempaa tarkastelua varten toteutettiin osittaja (Kuva 32, kohta 2) ja numero taulukko (Kuva 32 kohta 3).



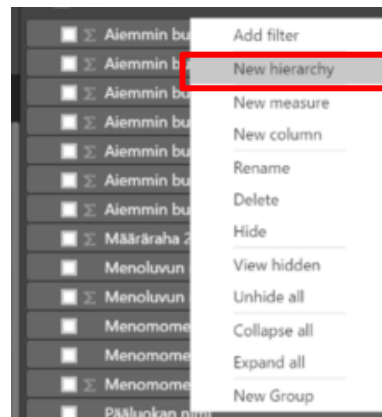
Kuva 32. Vesiputouskaavio, taulukko ja osittaja vuosien 2016–2017 menojen prosenttimuutoksesta.(Kuvakaappaus.)

## 8.2 Hierarkianäkymän toteutus

Talousarvion menojen ja tulojen hienojakoisemman jakaantumisen seuraamista yhdellä kaavionäytöllä helpottaa hierarkisen kaavion käyttö. Hierarkia auttaa päätelmien teossa, kun on mahdollista keskittyä vain yhdessä kaaviossa tapahtuvaan analyysiin. Talousarvion menojen hienojakoisemmassa seurannassa voidaan käyttää esimerkiksi hierarkian päälle rakennettua vaakapylväskaaviota (Kuva 37).

Hierarkisen vaakapylväskaavion rakentaminen aloitettiin kaaviotyyppin valinnalla. Tämän jälkeen valittiin Fields-paneelistä kenttä "Pääloukan nimi", jonka alle hierarkinen rakenne luotiin. Pääloukan nimi -kentässä hiiren kakkospainikkeen takaa perustettiin uusi hierarkia toiminnolla "New hierarchy" (Kuva 33). Toimenpiteen jälkeen muodostui uusi hierarkiataason ryhmä, jonka ensimmäisenä ryhmittelevänä tekijänä oli "Pääloukan nimi" (Kuva 34). "Pääloukan nimen" alemmalle tasolle siirrettiin vetämällä muut ryhmittelevät kentät "Menoluvun nimi" ja "Menomomentin nimi" (Kuva 35). Tämän jälkeen hierarkinen rakenne oli valmis.





Kuva 33. Uuden hierarkian luonti. (Kuvakaappaus.)

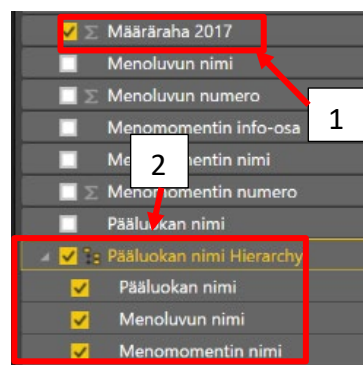


Kuva 34. Hierarkia ryhmä, jonka päätasolla on "Pääloukan nimi". (Kuvakaappaus.)



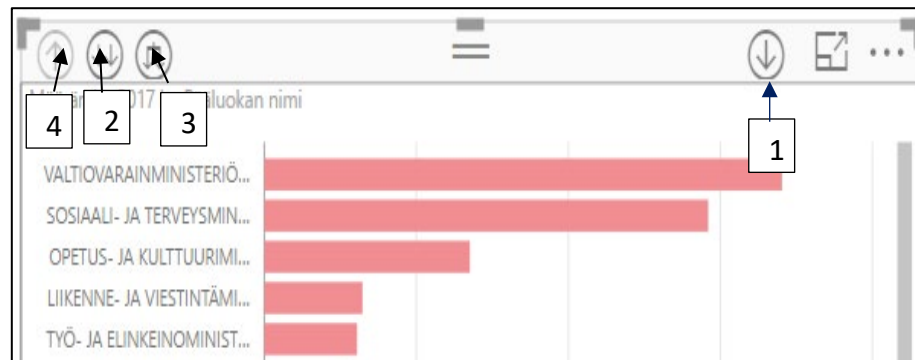
Kuva 35. Pääloukan alle tulevat alahierarkiat. (Kuvakaappaus.)

Hierarkian varsinaisessa käytössä valittiin arvo, jonka mukaan hierarkiaa haluttiin vaakapylväskaaviossa tarkastella. Arvoa määritteleväksi tekijäksi valittiin kenttä "Vuoden 2017 Määrärahat" (Kuva 36, kohta 1). Määrärahojen seurannassa seuranta tehtiin kaikilla hierarkiatasoilla, joten kentiksi valitaan "Pääloukan nimi", "Menoluvun nimi" ja "Menomomentin nimi" (Kuva 36, kohta 2). Tämän valinnan jälkeen hierarkioiden mukaan tarkasteltava vaakapylväs kaavio oli valmis.



Kuva 36. Kaavion valittavat hierarkiat ja tarkasteltava arvo "Määräraha 2017". (Kuvakaappaus.)

Kaavion analysoinnissa porautuminen esimerkiksi Sosiaali- ja terveysministeriön menoihin menoluvun ja menomomentin mukaan aloitettiin valitsemalla oikeassa reunassa olevan Drill Down -nuoli On-tilaan (Kuva 37, kohta 1). Hierarkiassa pääsee hienojakoisemmalle tasolle kaksoisnuolella (kuva 37, kohta 2) ja hierarkian voi avata kokonaan yhteneväisellä kaksoisnuolella (kuva 37, kohta 3). Takaisin ylätasolle päästään yhdellä nuolipainikkeella (kuva 37, kohta 4).



Kuva 37. Eri hierkiatasoilla tarkasteltava kaavio. (Täydennetty kuvakaappaus.)

### 8.3 Kustomoidut visualisoinnit

Microsoftin yhteistyökumppanit tarjoavat kustomoituja visualisointeja, jotka on ladattava erikseen Power BI:hin Office Storen -internet-sivuilta. Tämän luvun tavoitteena on tuoda esiin vaihtoehtoisia visualisointeja talousarvion seurantaan. Tarkoituksena on tuoda uusia näkökulmia ja tapoja esittää informaatiota, puu- ja verkostomallin, sekä yhdistelmägrafiikan avulla. Uusia esityskäytännöitä, kuten tässä luvussa esiteltäviä visualisointeja, käytetään talousaineistojen esittelyssä yhä enemmän. Käytettäessä uutta esitystapaa, tulee miettiä niiden soveltuvuutta käyttötarkoitustaan varten.

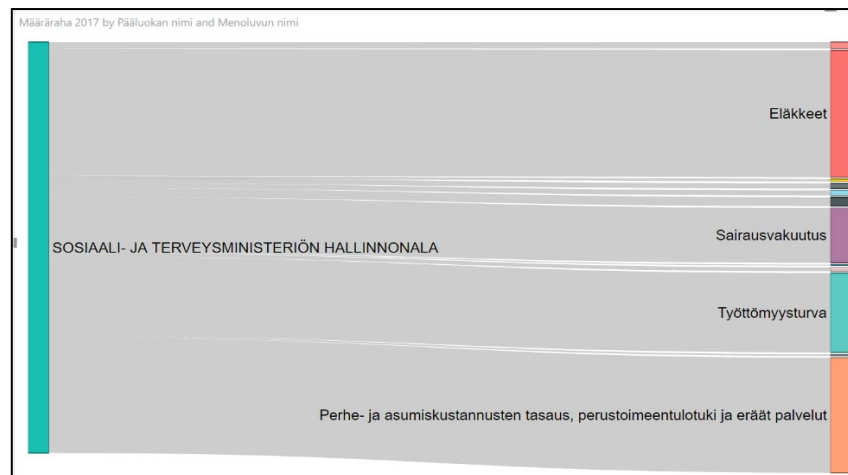
Luvussa ei käydä läpi, miten nämä uudet esittämistavat on rakennettu. Periaatteessa rakentaminen sujuu samalla tavalla kuin perus kaavioiden tekeminen. Poikkeuksena on infografiikan avulla luodut yhdistelmägrafiikat, joissa on hiukan enemmän työvaiheita normaalin kaavion rakentamiseen verrattuna.

#### 8.3.1 Puutyypinen virtauskaavio

Virtauskaavio eli Sankey-kaavio, muistuttaa puumaista rakennetta. Kaaviolla tarkastellaan Sosiaali- ja terveysministeriölle vuodelle 2017 budjetoituja menomäärärahoja menolukujen mukaan (kuva 38).

Shankey-kaaviossa viivojen paksuudella kuvataan yhteyksien voimakkuutta. Kuviotyypillä jaetaan kokonaisuudet hierakisiin osiin. (Koponen

ym. 2016, 236.) Kuvasta näkyy esimerkiksi se, että menomäärärahoja on budjetoitu eniten eläkkeisiin. Yhteyden voimakkuus Sosiaali- ja terveysministeriöön on vahva. Lähes yhtä vahvoja ovat viivat, sairausvakuutus, työttömyysturva sekä perhe- ja asumiskustannusten, perustoimeentulon menomäärärahojen osalta.

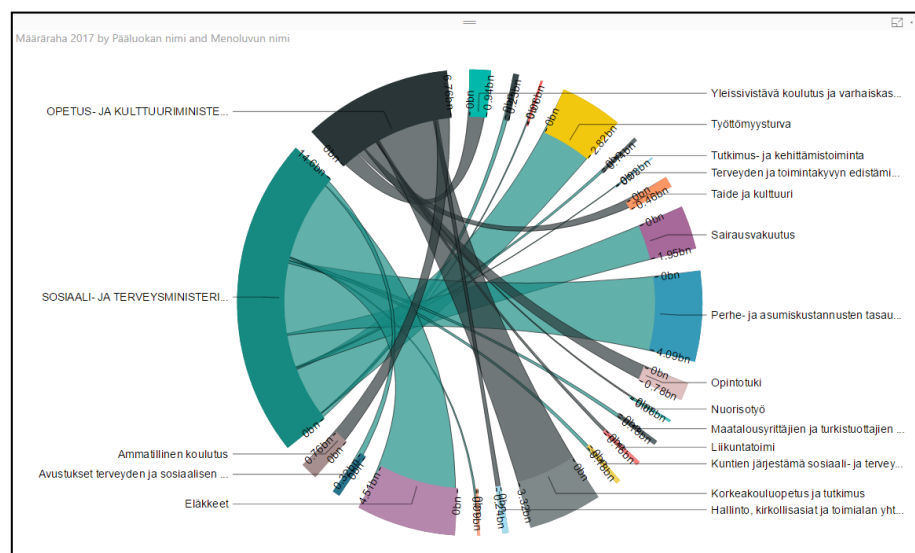


Kuva 38. Shankey-kaavio Sosiaali- ja terveysministeriön menoista menoluvun mukaan. (Kuvakaappaus.)

### 8.3.2 Verkostomalli

Verkostomallilla voidaan kuvata esimerkiksi kohteiden välille rakennetuilla kaarilla. Kaaret kuvaavat kohteiden välistä yhteyttä. Kaaret voivat olla saman vahvuisia tai vahvuus voi vaihdella riippuen siitä kuinka vahva yhteys kohteiden välillä on. (Koponen ym. 2016, 231.)

Voimaohjatussa jännekuviassa (Chord) tarkasteltiin Sosiaali- ja terveysministeriölle, sekä Opetus- ja kulttuuriministeriölle vuodelle 2017 budjetoituja menomäärärahoja menoluvuittain. Kuvan voimaviivojen perusteella voidaan edelleen nähdä, että suurimmat menoerät kohdistuvat eläkkeisiin, sairausvakuutus, työttömyysturva, sekä perhe- ja asumiskustannusten menoihin. Lisäksi Opetus- ja kulttuuriministeriön kohdalla suurimmat budjetoidut menomäärärahat kohdentuvat korkeakouluopetukseen ja tutkimukseen (Kuva 39).



Kuva 39. Chord-kaavio Sosiaali- ja terveysministeriön, sekä Opetusministeriön budjettimäärärahoista menoluvuittain. (Kuvakaappaus.)

### 8.3.3 Yhdistelmägraafiikka

Yhdistelmägraafiikka on moniulotteinen esitystapa, jossa on yhdistetty esimerkiksi tietokuvitusta, tilastografiikka ja tekstiä rinnakkain. Yhdistelmägraafiikasta käytetään monissa yhteyksissä käsitettä *infografiikka*. Yhdistelmägraafiikka terminä ei ole yksiselitteinen, koska esitystavassa on käytetty useita visuaalisen esittämisen elementtejä. (Koponen ym. 2017, 141.)

Alla olevassa kuvioissa on yhdistelmägraafiikan avulla esitetty budjettimäärärahojen jakaantumista pääluokittain (Kuva 40). Kriteerinä oli, että budjettimäärärahan tulee olla vähintään 1 mrd. euroa. Kuvasta on nähtävissä, että määrärahoja on budjetoitu eniten Valtiovarainministeriölle. Lisäksi kuvioissa esitetään muiden pääluokkien suhteellinen osuus VM:n määrärahoista. Kuviota voi myös tarkastella osittajan avulla.



Kuva 40. Yhdistelmägraafiikka, jossa budjettimäärärahojen jakaantuminen on esitetty pääluokittain. (Kuvakaappaus.)

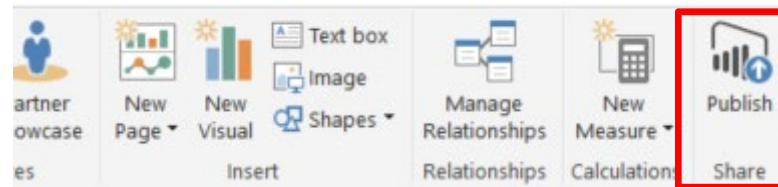
#### 8.4 Talousarvion raporttien julkaiseminen Power BI -pilvipalvelussa

Ennustetoimistossa on myös tarvetta jakaa raporttiedostoja käyttäjien kesken, toimistossa tai koko pankissa. Raporttien jakaminen tapahtuu organisaatiokohtaisten ryhmien kautta. Ryhmille on mahdollista määritellä erilaisia käyttöoikeuksia. Lisäksi raportteja jaetaan sosiaalisessa mediassa, tarvittaessa niitä halutaan upottaa myös internet-sivustoille. Raporttiedostoista on mahdollista toteuttaa koontinäytöjä, jotka näkyvät vain Office365-palvelun käyttäjille. Kaikki nämä toiminnot tapahtuvat Power BI:n Service -palvelussa, Power BI:n pilvipalvelun julkaisualustalla.

Tässä luvussa käydään läpi miten Power BI Desktopilla luotu talousarviotiedosto, joka koostuu raporttinäkymistä ja dataseteistä, sekä tietomallista, julkaistiin Power BI:n pilvipalvelussa pbix-tiedostona. Tämä tarkoittaa käytännössä sitä, että tiedosto siirretään Power BI:n pilvipalvelussa työtilaan (Workspace). Tiedosto, tietomalli ja raportit siirtyvät työtilaan samalla nimellä kuin ne olivat Desktopissa. Power BI -pilvipalvelussa on vielä mahdollista tehdä muokkauksia esimerkiksi kaavioihin. Pilvipalvelun tarjoamat mahdollisuudet kaavioiden ja aineistojen käsittelyyn ovat huomattavasti rajoittuneemmat kuin Desktopissa. Huomioitavaa on myös se, että kaikki muutokset, jotka tehdään Power BI:n pilvipalvelussa, eivät enää tallennu Desktopiin. (powerbi.microsoft.com; hexcelligent.fi 2016.)

Desktopissa tiedosto julkaistiin Power BI Desktopin valintanauhalla Julkaise (Publishing) -valinnalla (Kuva 41). Power BI -pilvipalveluun kirjaututaan esimerkiksi työpaikan sähköpostin Office 365-käyttäjätunnuksella, jonka kautta tunnukset Power BI -pilvipalveluun on luotu. Tässä työssä

Office 365-pilvipalvelualustana käytettiin Hämeenlinnan ammattikorkeakoulun pilvipalvelua.



Kuva 41. Raporttien julkaiseminen. (Täydennetty kuvakaappaus.)

## 8.5 Talousarvion koontinäytön teko Power BI -pilvipalvelussa

Talousarvion seurannassa koontinäytön tavoitteena on antaa tilannekuva talousarvion pääkohdista. Näyttöön kerättiin visualisointeja talousarviosta ja lisätalousarvioista. Kyseessä oli siis kaksi eri tiedostoa, joista koottiin koontinäyttöön tärkeimmät seurattavat raportit.

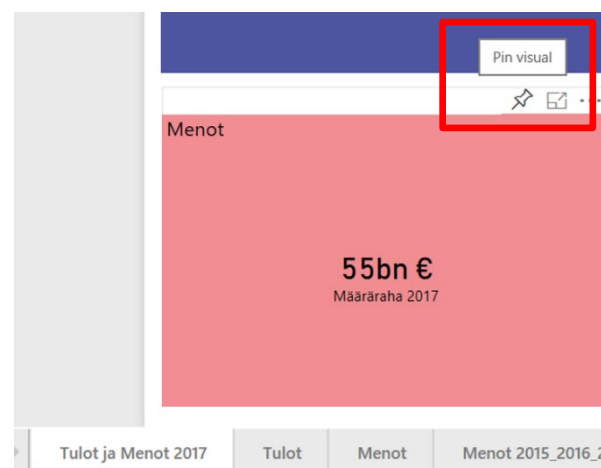
Power BI:n pilvipalvelussa on mahdollista kiinnittää raporttinäyttöjen visualisointeja valikoidusti. Koontinäyttöön voi valita oleellimmat visuaaliset näytöt omaa seurantaan tai esimerkiksi organisaation erityistä kohderyhmää varten. Koontinäytön luonti aloitettiin valitsemalla pilvipalvelun vasemmasta paneelista valinta "Dashboards" ja oikeasta reunasta + -merkki (Kuva 42, kohta 1).

Koontinäytölle annettiin nimeksi "Talousarvion seuranta" (Kuva 42, kohta 2). Tämän jälkeen kerättiin koontinäytölle tarvittavat raportit. Ensimmäinen siirryttiin TalousarvioV4-tiedoston raporttinäytölle jossa valittiin sopivat kaaviot koontinäyttöön (Kuva 42, kohta 3).



Kuva 42. Uuden Dashboardin luonti ja raportteja sisältävän TalousarvioV4-tiedoston valinta. (Täydennetty kuvakaappaus.)

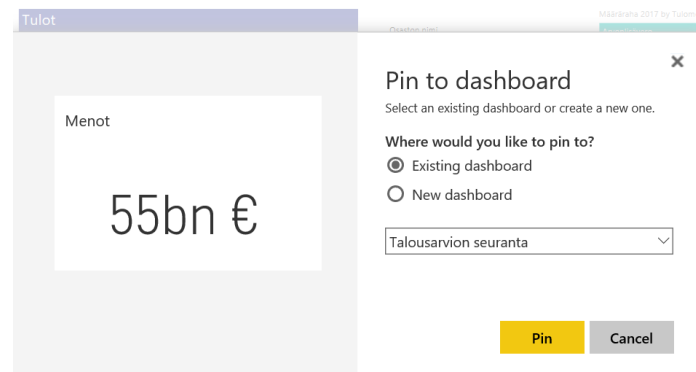
Koontinäytölle kiinnitettiin TalousarvioV4-tiedostosta visualisointi nimeltä "Menot" kaavion oikeasta yläkulmasta Pin visual -kuvakella (Kuva 43).



Kuva 43. Kaavion kiinnittäminen koontinäyttöön Pin visual -toiminnolla. (Täydennetty kuvakaappaus.)

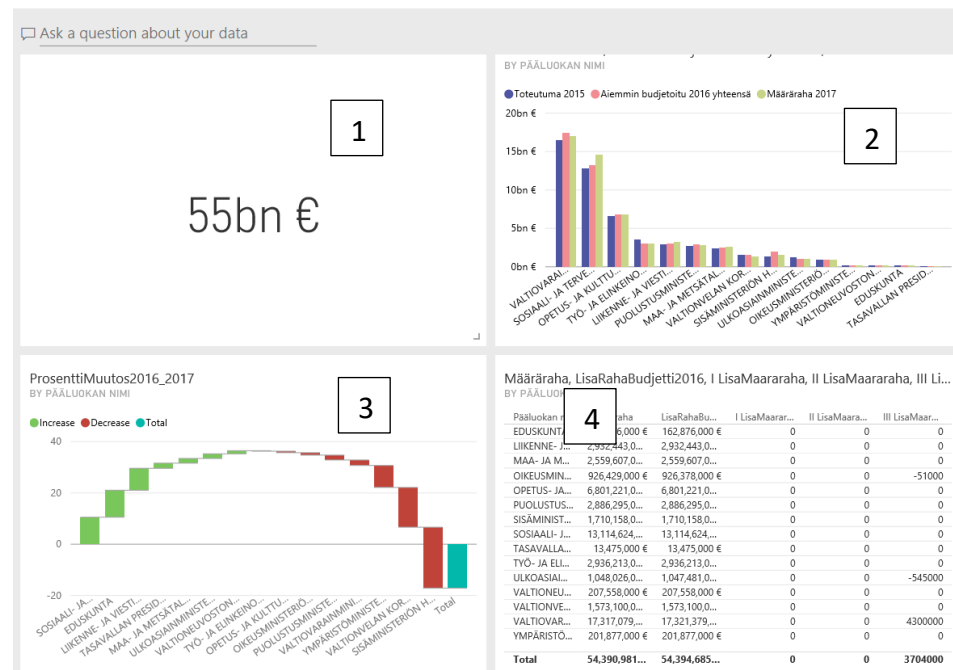
Uusi koontinäyttö perustettiin oletusnimelle "Talousarvion seuranta". Oletusnimi hyväksyttiin Pin-toiminnolla (Kuva 44). Tämän jälkeen kaavio oli tarkasteltavissa myös koontinäytöllä (Kuva 45, kohta 1). Koonti-

näytölle lisättiin vielä samasta raportti-tiedostosta visualisointi menoista vuosilta 2015–2017 (Kuva 45, kohta 2), ja vesiputouskaavio prosenttimuutoksista (Kuva 45, kohta 3).



Kuva 44. Raporttien valinta ja kiinnittäminen koantinäytölle. (Kuvakaappaus.)

TalousarvioLisatalousarvioV2-tiedostosta valittiin vielä taulukko vuodelle 2016 budjetoiduista lisämäärärahoista. Taulukko siirrettiin myös koantinäytölle (Kuva 45, kohta 4). Koantinäyttölle kerätyt visualisoinnit mahdollistavat nyt kokonaisvaltaisemman menoseurannan, joka ei edellytä yksittäisten raportti-tiedostojen avaamista (Kuva 45).



Kuva 45. Koantinäyttö vuoden 2017 budjetoiduista kokonaismenoista, 2015-2017 budjettivertailuista, vuoden 2016–2017 prosenttimuutoksista. (Kuvakaappaus.)



## 9 KOKEMUKSET POWER BI:N KÄYTÖSTÄ

Power BI:n käyttö on perustyöskentelyn näkökulmasta kohtuullisen helppoa. Osa välineen perustoiminnoista vastaa taulukkolaskentaohjelma Excelin toimintoja. Exceliä käyttäneille Power BI:n käyttöönottoon onkin todennäköisesti matala kynnyks. Ohjelman monipuolisempaan käyttöön ja hyödyntämiseen vaaditaan kuitenkin syvällisempää tietojenkäsittely- ja ohjelmointiosaamista. Ohjelman kaikkien ominaisuuksien käyttöönotto vaatii todennäköisesti Power BI koulutusta, sekä myös ohjelmointikielten opiskelua.

Toisaalta kuten Excelinkin käyttö on aikojen saatossa vaatinut monia toistokertoja, sama pätee varmasti Power BI:n käyttöön. Tutustumalla ja testaamalla välinettä aktiivisesti, saadaan paras mahdollinen mielikuva Power BI:n soveltuvuudesta ja sen tuomista mahdollisuuksista mm. Suomen Pankin toimintaympäristössä.

### 9.1 Taulujen käsittely

Power BI:lla haastavinta oli talousarviodatan tuomisen automatisoinnin toteuttaminen. Automatisointi vaatii edelleen tutkimista, niin että datan hyödyntäminen ulkoisesta tietolähteestä olisi joustavaa. Lisäksi aineiston siivoaminen käyttökelpoiseen kuntoon vaatii panostusta. Talousarvion osalta data oli kuitenkin laadultaan hyvää, joten datan puhdistaminen jäi melko vähäiseksi datan tuontivaiheessa.

Haasteita talousarviodatan käsittelyssä aiheutti lähteestä haettavien datataulujen suuri määrä, sekä Power BI:n oletusarvoisesti muodostamat tietomallin taulujen väliset yhteydet. Tätä työtä tosin helpotti taulujen yhdistäminen Append Query -toiminnolla. Power BI:n tuottamat tietomallit onkin hyvä tarkastaa datan käyttöönotto vaiheessa.

Taulujen sisällön käsittelyn liittyvien toimintojen osalta oli vaikeaa hahmottaa, mitä toimintoja oli mahdollista tehdä Data-näkymän Home- ja Modeling -välilehdellä. Erityisesti Home-välilehden Edit Queries -toiminnon ja Modeling-välilehden toimintojen väliset yhteydet välillä hämärtyivät, mm. laskennan ja tietotyyppimuunnosten osalta. Taulujen käsittelyn tulisikin olla selkeämpää. Pääsääntöisesti Home-välilehden Edit Query -toiminnot ovat kattavammat ja ne koskevat enemmän taulujen rivien ja sarakkeiden käsittelyä, kun taas Modelling-välilehdellä käsitellään tietomallia, laskentaa, sarakkeiden järjestelyä ja tietotyyppien muokkausta.

## 9.2 Laskenta

Laskentakaavojen teko on periaatteessa helppoa, jos DAX-funktioiden käyttö on tuttua. DAX-funktioita on ollut mahdollista käyttää myös Excelin Power Queryssä. DAX-laskenta voi olla haasteellista niille, jotka eivät ole tottuneet käyttämään sarakkeiden nimiä ja kenttiä laskennassa, jota perinteinen Excelillä laskeminen ei edellytetä.

Joissakin tapauksissa monimutkaisten laskelmien tekeminen ja If-funktioiden käyttö on helpompaa Excelissä (Laksola ym. 2016, 34). DAX-funktiot tarjoavat kuitenkin paljon monipuolisia mahdollisuuksia laskennan suorittamiseen, joten funktioiden käyttöön ja niiden tarjoamiin mahdollisuuksiin on hyvä tutustua. Opinnäytetyön teon aikana myös selvisi että Report-näkymän Quick Measure -toiminto pitää sisällään monia hyödyllisiä funktioita, joita käyttämällä voidaan korvata monet mutkikkaat DAX-laskukaavat.

## 9.3 Visualisointi

Visualisointi oli hyvin helppoa ja kaavioiden vaihtaminen nopeaa. Kaavioiden muokkaus esimerkiksi otsikoiden, fonttien X- ja Y-akseleiden, sekä värimäärittelyjen osalta, onnistuu sujuvasti. Lisäksi muokkauksmäärittelyt löytyvät helposti visualisointipaneelistä. Eri muuttujia on helppo valita kaavioihin nopeasti. Power BI:hin voidaan lisätä organisaation omia teemoja, joissa on määritelty esimerkiksi organisaation oma väri- ja fontti-maailma. Tämä on hyödyllinen ominaisuus, jota Suomen Pankissakin varmasti käytettäisiin. Visualisoinnin helppo toteuttaminen on Power BI -välineen vahvuuksia.

Power BI:n tarjoamia kaaviotyyppejä on melko vähän tarjolla, tosin tarjolla on erikseen ladattavia kustomoituja kaavioita. Näistä tarjontaa löytyy tilastollisista kaaviotyypeistä laajemminkin, sekä lisäksi myös esimerkiksi infografiikasta. Ohjelmisto-osaajat ovat laatineet kustomoidut visualisoinnit Power BI:n verkkosivustoille, joten Microsoft ei ota vastuuta niiden toimivuudesta. Kustomoitujen kaavioiden käyttö oli yhtä helppoa kuin normaalien peruskaavioiden käyttö. Infografiikan luominen pitää sisällään enemmän työvaiheita ja oli hiukan haastavampaa.

Kaavioiden eli ns. raporttien, sekä koontinäyttöjen liittäminen esimerkiksi Power Pointiin tapahtuu helpoiten pilvipalvelussa Export-toiminnolla. Kaavioita ei voi linkittää Power Pointiin. Suomen Pankissa käytetään paljon Power Point-esityksiä, joten linkitysten teko Power Pointiin on tärkeä ominaisuus, jonka pitäisi toimia. Kuinka paljon sitten Power BI:n kautta julkaistavat esitykset voisivat korvata Power Point-esitykset, on oma kysymyksensä. Tämän hetken tilanteen osalta uskoisin,

että Power Point-esitysten teko on hyvin vakiintuneessa asemassa, joten sen korvaavan välineen pitäisi tuoda paljon lisäarvoa nykyiseen välineeseen verrattuna.

#### 9.4 Pilvipalvelu

Pilvipalvelun käyttäjäryhmien hallinnointia ei tässä työssä päästy konkreettisesti testaamaan, joten sen käytön arviointia ei pystytty tekemään. Omien aineistojen kuten raporttien ja koontinäyttöjen hallinnointi sen sijaan on helppoa pilvipalvelussa.

Suomen Pankissa korostetaan tiedon hallintaan ja etsimiseen liittyviä ominaisuuksia ja näitä ei vielä tähän mennessä ole Power BI Servicessä tarjolla. Hallinta keskittyy lähinnä ryhmäkohtaiseen hallintaan.

#### 9.5 Koontinäyttö

Koontinäytön toteutuksen tekeminen oli erittäin helppoa. Koontinäytön käytössä hämmennystä saattaa herättää se että kaaviota aktivoimalla näkymä siirtyykin suoraan Raportti-näkymään. Paluu koontinäytölle on kuitenkin helppoa. Lisäksi monelta käyttäjältä saattaa jäädä huomaamatta koontinäytön tarjoama Insight-mahdollisuus, joka johdattaa käyttäjän yksittäisen kaavion monipuolisempaan tarkastelunäkymään. Insight-näkymässä on käytetty nopeaa oivallusta, jonka avulla ohjelman koneäly tekee nopeita dataan liittyviä päätelmiä. (powwerbi.microsoft.com; Laksola ym. 2016, 33.)

Koontinäytön jakamisen rajoittaminen vain Office 365-käyttäjille ei välttämättä selviä käyttäjille heti. Koontinäytön käytön rajoittaminen on perusteltua, koska se mahdollistaa siirtymisen raporttinäkymään jossa data sijaitsee, sitä ei haluta välttämättä mahdollistaa ulkopuolisille käyttäjille. Koontinäytön kaltaisen näkymän, jonka voi jakaa internetissä, on mahdollista rakentaa Raportti-näkymien avulla.

## 10 POWER BI:N HAASTEET

Power BI on loppukäyttäjän käytettäväksi tarkoitettu itsepalvelu eli ns. Self Service- ja pilvipainotteinen väline. Välineellä voidaan nopeasti jalostaa tiedosta esitysaineistoa esimerkiksi Ad Hoc- ja analyysikäyttöön.

Power BI-välineenä tuo datan analysointiin, jalostamiseen ja visuaaliseen esittämiseen mahdollisuuksia mutta myös haasteita. Power BI:n datan käsittelyä tulisi vielä selkeyttää, jotta sen käyttäjäkunta kasvaisi ja Power BI:ta voitaisiin ryhtyä kutsumaan täysveriseksi Self Service-välineeksi.

### 10.1 Datan käsittelyn ja kuviohallinnan haasteet

Power BI ei ole ns. ETL-väline, sillä ei voi päivittää tai ylläpitää esimerkiksi aikasarjatietokannan tietoja. Näin ollen mm. ennustetoimiston datan käsittelyn työvaiheisiin liittyen välinettä ei voida pitää yhdenlukuun välineenä, vaan datan ylläpidossa joudutaan käyttämään toisia välineitä, kuten Exceliä.

Power BI ei tarjoa kuviohallinnan ja aineistojen hakutoimintoihin liittyviä palveluita. Näin ollen Power BI on pikemminkin väline, joka tukee visuaalista esittämistä muiden käytössä olevien välineiden rinnalla. Monien ohjelmointikielien käyttömahdollisuus antaa mahdollisuuksia, mutta tuo myös haasteita.

Välineiden osalta tulisi osata valita oikea kieli, jota olisi tehokkainta hyödyntää halutun lopputuloksen saamiseksi. Power BI:n tehokas ja syvällisempi käyttö kaiken mahdollisen hyödyn saamiseksi, vaatii harjaantunutta dataosaamista sekä ohjelmointitaitoja.

### 10.2 Pilvipalvelun haasteet

Data säilötään Microsoftin mukaan pilvipalveluissa EU-alueen sisäpuolelle, joten data on EU-alueen lainsäädännön piirissä. Koska kyseessä on pilvipalvelu, tulisi palvelua hankittaessa selvittää, mitä palvelu kattaa, mitkä ovat sen vasteajat ja mitkä ovat toimittajan vastuut.

Microsoft tarjoaa Premium-versiossaan pilvipalveluiden osalta myös hybridivaihtoehdon, jonka tarjoamia mahdollisuuksia olisi hyvä tutkia. Eri palvelupakettien, kuten Power BI Pron, ja Premiumin valintojen välillä tulisikin osata tehdä oikeita päätöksiä, laajempaa käyttöä harkittaessa.

### 10.3 Käyttäjätuki

Microsoftin antama tuki välineen käytölle on keskittynyt Power BI:n yhteisöjen antamaan tukeen. Käyttäjätuen saaminen suoraan Microsoftin asiakastuesta saattaakin olla haasteellista. (Laksola ym. 2016, 25.)

Lisäksi erilaiset sisältöpaketit ja erilaisten käyttöä helpottavien kokonaisuuksien tarjonta hoidetaan yhteistyökumppaneiden kautta. Käytännössä tuen saanti näihin paketteihin kokonaisuuksiin on täysin yhteistyökumppaneiden varassa. Microsoft ei ota vastuuta niiden päivityksestä tai ylläpidosta.

## 11 TULOKSET

Lähtötilanteeseen verrattuna uuden talousarviodatan käyttöönotto tuo uusia mahdollisuuksia datan laajempaan hyödyntämiseen. Seurannan ja ennustamisen tueksi rakennetut raporttinäytöt sekä koontinäyttö vastasivat niille asetettuja vaatimuksia. Välineen käytön osalta datan käsittely oli haasteellista.

### 11.1 Tulosten arviointi

Avoimen talousarviodatan hyödyntäminen Power BI:lla tuo paljon lisäarvoa datan käyttöön verrattaessa sitä lähtötilanteeseen, jossa dataa päivitettiin käsin. Käytettävän aineiston määrä on huomattavasti suurempi, informaatiota on saatavilla karkealla, mutta myös yksityiskohtaisemmalla tasolla. Näin ollen mahdollisuudet datan syvällisempään pureutumiseen ja analysointiin ovat monipuolisemmat. Uusien mahdollisuuksien myötä ennuste- ja seurantatyöhön saadaan talousarvion käytön osalta perusteellisempaa tietoa, joka auttaa täsmällisempien johtopäätösten teossa.

Uudet raporttinäytöt helpottavat asiantuntijoiden työtä, mm. silloin kun aineiston osalta on tehtävä nopeita analyyskejä ja johtopäätöksiä. Valmiiksi määritellyjä visualisointeja on helppo muokata nopeasti omien tarpeiden mukaan. Lisäksi visualisoinnit tarjoavat mahdollisuuksia datan tarkasteluun useamman kaavion välillä, tai esimerkiksi hierarkiatasolla suoraan kaavion sisällä. Visualisointeja talousarviodatasta ei aikaisemmin ole ollut käytössä, joten perusraportit tehostavat huomattavasti asiantuntijan työtä. Lisäksi yhdelle raporttinäytölle koostetut kaaviot on helppo jakaa internetissä, ja niitä voidaan käyttää myös esittelyaineistoina Suomen Pankin ulkopuolisille käyttäjäkunnille. Tilanteen mukaan raportteja voidaan nopeasti upottaa myös Twitteriin tai muihin sosiaalisiin medioihin.

Koontinäyttö tuo uusia mahdollisuuksia, kun esimerkiksi kuluvan vuoden budjetoituja määrärahoja ja lisätalousarvioita voidaan seurata kootusti yhdellä ja samalla koontinäytöllä. Eri Power BI-tiedostoissa olevat visualisoinnit, tai kokonaiset raporttinäytöt on todella helppoa yhdistää yhdelle koontinäytölle vaivattomasti. Lisäksi ns. Insight-näkymä tuo uutta lisäarvoa koontinäytön yksittäisten kaavioiden syvällisempään analysointiin. Power BI:n sisäänrakennettu niin sanottu koneäly osaa automaattisesti rakentaa visualisointien pohjalta syvempää analyysiä käyttäjän tueksi. Koontinäytöllä Power BI:n Insight-näkymässä raporttinäyttökohtainen automaattinen analysointi helpottaa asiantuntijan työtä. Koontinäytöillä voidaan helposti myös tuoda uutta ulottuvuutta pankin sisäisissä esityksissä johtokunnalle ja koko henkilökunnalle.

## 11.2 Jatkokehityshankkeet

Power BI tarjoaa paljon mahdollisuuksia ja ominaisuuksia, joita tämän opinnäytetyön aikana ei ollut mahdollisuutta testata. Näistä ominaisuuksista olisi hyvä muodostaa konkreettinen tuntuma. Ennustetoimistossa on parhaillaan käynnissä Pro-lisenssien hankinta. Tämä mahdollistaa tutustumisen Power BI:n ominaisuuksiin laajemminkin. Power BI:n tarjoamia mahdollisuuksia tulisi pohtia ennustetyön näkökulmasta, niiden tärkeysasteen mukaan. Tämän jälkeen tärkeimpien ominaisuuksien testaaminen tulisi priorisoida jatkokehityshankkeisiin.

Lopputyön loppuvaiheessa saatiin lisäinfoa XML-muotoisen talousarviodatan automatisoinnista. Automatisoinnin toteutukseen ja suunnitteluun on kuitenkin käytettävä aikaa, joten se toteuttaminen tulisi siirtää mahdollisiin jatkokehityshankkeisiin. Lisäksi datan yhdistelyä toisiin datalähteisiin kuten tietokannassa olevien julkisen talouden aikasarjoihin kannattaa testata. Power BI:n käyttöä tulisi laajentaa muunkin ennustetoimiston ennustetyön aineiston käytön testaamiseen. Tietolähteinä on käytettävissä noin 70 eri lähdettä, joten Power BI saattaa tarjota mahdollisuuksia myös ekonometristen ohjelmien tuottamaan datan käsittelyyn.

Power BI tarjoaa analyysivälineitä, jotka saattaisivat tuoda lisäarvoa ennuste- ja seurantatyöhön. Oletuksena analyysi-paneelissa on tarjolla dynaamiset viiteviivat minimi, maksimi, keskiarvo, prosentti- ja mediaani arvoille. Analytics pane-osioon on myös mahdollista asentaa Forecasting-lisäosa.

Power BI mahdollistaa monien ohjelmointikielten, kuten DAX, M, SQL ja R-ohjelmointikielten käytön. Tilastollinen R-ohjelmointikieli on yksi osa-alue, joka saattaisi tuoda paljonkin uusia mahdollisuuksia, automaattiseen tiedonsiirtoon, laskentaan, sekä visuaalisten näyttöjen toteuttamiseen ja hyödyntämiseen.

Ryhmätyötilojen käyttö ja aineistojen sekä visuaalisten näyttöjen jakaminen on mahdollista Power BI Pro -lisenssien käyttäjille. Ryhmätyökentelyä sekä raporttien ja koontinäyttöjen jakamista tulisikin testata aktiivisesti lisenssin käyttöönoton yhteydessä. Ryhmätyötilojen käytön toimintaperiaatteisiin ja logiikkaan on hyvä perehtyä syvällisesti. Näin voidaan saavuttaa järkevät käytännöt tiedon hallinnan näkökulmasta.

Tämän lisäksi kannattaisi tutustua Rest-rajapintoihin ja niiden yhteyksiin Azure-palveluun. Power BI:n tarjoamien mahdollisuuksiin datan louhinnassa R-kielellä. Näiden osa-alueiden osalta tulisikin toimiston tietämystä ja osaamista kasvattaa.

## 12 YHTEENVETO

Käytännön osuutta arvioitaessa monipuolisempi käytössä oleva aineisto tuo syvyyttä seuranta- ja ennustetyöhön. Raportti- ja koontinäytöt tuovat lisää mahdollisuuksia datan analysointiin, seurantaan ja nopeaan työskentelyyn. Näin ollen tuotetut perusraportit ja koontinäyttö täyttävät hyvin empiirisen osuuden tavoitteet ja vastasivat kysymykseen "Miten Power BI:n avulla voidaan hyödyntää avoimen talousarviodatan käyttöä seuranta ja ennustetyössä?".

Avoimen talousarviodatan käyttöönoton avulla voidaan arvioida säästettävän aikaa ja kustannuksia, kun datan päivitykseen liittyvät manuaaliset työvaiheet vähenevät. Työ tehostuu, kun analysoidua ja jalostettua aineistoa tuotetaan eri käyttötarkoituksia varten nopeasti. Datan käyttöönoton toteutus sekä raporttien ja koontinäytön rakentaminen vaativat työaikaa. Lisäksi uuden välineen käytön opettelu ja sen tarjoamiin mahdollisuuksiin tutustuminen vaatii oman työpanoksensa. Näin ollen avoimen talousarviodatan käyttöönoton ja jatkojalostamisen kehitystyöstä syntyi kustannuksia. Datan käytöllä saavutettavia taloudellisia vaikutuksia on vielä tässä vaiheessa vaikea arvioida.

Tutkimuskysymykseen "Mitkä ovat avoimen datan taloudelliset vaikutukset?" ei pystytä antamaan täsmällistä vastausta. Taloudellisten vaikutusten mittaamiseen ja niiden arviointiin ei tähän mennessä ole vielä löydetty oikeita välineitä ja menetelmiä. Mittaamisen ja menetelmien kehittämisestä ollaan kiinnostuneita ja niihin panostetaan voimakkaasti. Todennäköisesti tulevaisuudessa avoimen datan taloudellisesta vaikuttavuudesta on saatavissa tietoa, joka perustuu täsmällisiin tutkimustuloksiin.

Opinnäytetyön tekeminen oli haastavaa mutta opettavaista. Työ tehtiin hyvin pitkälle itsenäisesti oman työn ohessa. Avoimen datan teoreettisen osuuden soveltaminen käytäntöön vaati pohdintaa niiden keskinäisten yhteyksien ymmärtämisessä. Työn tavoitteena oli herättää lukija pohtimaan, miten avointa dataa voidaan hyödyntää, ja osoittaa konkreettisten tulosten pohjalta, että datan käyttö on kannattavaa. Rohkaisisinkin organisaatioita ja yrityksiä pohtimaan sitä, miten avointa dataa kannattaisi hyödyntää omassa toiminnassa - käyttö luo mahdollisuuksia. Työn aikana tietämykseni Power BI:n käytöstä välineenä kasvoi valtavasti. Valitettavasti kaikkea uutta tietoa ei tämän työn aikana pystytty käytännössä testaamaan. Kehitystyö jatkuu ja välineen käyttöön tullaan jatkossakin perehtymään Suomen Pankin ennustetoimistossa laajemman käyttäjäkunnan keskuudessa.



## LÄHTEET

Ansip, A. (2016). *Open data across the Atlantic*. Blogijulkaisu 15.11.2016. Haettu 6.3.2017 osoitteesta [https://ec.europa.eu/commission/commissioners/2014-2019/ansip/blog/open-data-across-atlantic\\_en](https://ec.europa.eu/commission/commissioners/2014-2019/ansip/blog/open-data-across-atlantic_en)

A University for the web P2PU (2.3.2017). Built by open community. *Datan avaaminen johdanto*. Haettu 3.3.2017 osoitteesta <https://courses.p2pu.org/en/courses/2486/datan-avaaminen-johdanto/>

Avoindata.fi 2015. *Valtion talousarvioesitykset*. Haettu 3.2 osoitteesta <https://www.avoindata.fi/data/fi/dataset/valtion-talousarvioesitykset>

budjetti.vm.fi n.d. *Talousarvioesityksen sisältö*. Haettu osoitteesta 3.2.2017 <http://budjetti.vm.fi/>

Bureau of Economic Analysis (BEA) (2016). *BEA Unveiling New Data Tool Aimed at Faster Access to Economic Statistics in United States, Europe*. Blogijulkaisu 8.11.2016. Haettu 4.8.2017 osoitteesta <https://blog.bea.gov/2016/11/08/bea-unveiling-new-data-tool-aimed-at-faster-access-to-economic-statistics-in-united-states-europe/>

Creative Commons.fi (2017). *Julkaise avoimesti*. Haettu 27.3.2017 osoitteesta <https://creativecommons.fi/valitse/>

Euroopan komissio (2015). *Creating Value through Open Data; Study on the Impact of Re-use of Public Data Resources*. Haettu 4.5.2017 osoitteesta [https://www.europeandataportal.eu/sites/default/files/edp\\_creating\\_value\\_through\\_open\\_data\\_0.pdf](https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf)

Gartner (2017). *Magic Quadrant for Business Intelligence and Analytics Platforms*. Haettu 14.4.2017 osoitteesta <https://www.gartner.com/doc/reprints?id=1-3TXXSLV&ct=170221&st=sb>

Helsinki Region Infoshare (2010). *Mitä avoin data on?* Haettu 1.3.2017 osoitteesta <http://www.hri.fi/fi/mita-on-avoin-data/>

HExcelligent.fi (2016). *Power BI - kaikki mitä sinun tulee tietää aloittaessasi*. Blogijulkaisu 30.1.2016. Haettu 15.3.2017 osoitteesta <https://hexcelligent.fi/2016/01/30/power-bi-kaikki-mita-sinun-tulee-tietaa-aloittaessasi/>

Hovi, A., Hervonen, H. & Koistinen, H. (2009). *Tietovarastot ja Business Intelligence*. Helsinki: WSOY.

Koponen, J., Hildén, J. & Vapaasalo, T. (2016). *Tieto näkyväksi: Informaatiomuotoilun perusteet*. Helsinki: Aalto-yliopiston taiteiden ja suunnittelun korkeakoulu, Aalto Arts Books.

Laksola, J. & Mauno, M. (2016). *Business Intelligence: Microsoft Power BI*. Kandidaatintyö. Tuotantotalous. Lappeenrannan teknillinen yliopisto. Haettu 5.4.2017 osoitteesta [https://www.doria.fi/bitstream/handle/10024/123487/Kandidaatintyo\\_Laksola\\_Mauno.pdf?sequence=2](https://www.doria.fi/bitstream/handle/10024/123487/Kandidaatintyo_Laksola_Mauno.pdf?sequence=2)

Microsoft (2017). *Guided Learning for Power BI*. Haettu 1.3.2017 osoitteesta <https://powerbi.microsoft.com/en-us/guided-learning/powerbi-learning-0-0-what-is-power-bi/>

Pelttari, S. (2014). *Avoin data*. Opinnäytetyö. Tietojenkäsittely. Oulun ammattikorkeakoulu. Haettu 27.2.2017 osoitteesta [https://www.theseus.fi/bitstream/handle/10024/86046/Pelttari\\_Sirpa.pdf?sequence=1](https://www.theseus.fi/bitstream/handle/10024/86046/Pelttari_Sirpa.pdf?sequence=1)

Romppainen, H. (2015). *Miksi avointa dataa? Monialue tutkimus Oulun elueen toimijoista*. Pro gradu -tutkielma. Tiedeviestinnän maisteriohjelma. Oulun yliopisto. Haettu 3.4.2017 osoitteesta <http://jultika.oulu.fi/files/nbnfioulu-201510152088.pdf>

Valtiovarainministeriö (VM) 31/2015. *Avoimesta datasta innovatiiviseen tiedon hyödyntämiseen*. Avoimen datan loppuraportti 2013 – 2015. Haettu 4.4.2017 osoitteesta <http://vm.fi/documents/10623/1107406/Avoimen+tiedon+ohjelman+loppuraportti/8eaaee68-6f3b-4a48-8b57-c5866315bf13?version=1.0>

Valtioneuvoston kanslia (VNK) 40/2017. *Avoimen datan hyödyntäminen ja vaikuttavuus*. Haettu 17.3.2017 osoitteesta [https://www.etla.fi/wp-content/uploads/VNK\\_2017\\_40.pdf](https://www.etla.fi/wp-content/uploads/VNK_2017_40.pdf)

Valtiovarainministeriö (VM) 15a/2015. *Avoimen tiedon vaikuttavuus - esitutkimus*. Haettu 26.3.2017 osoitteesta <http://vm.fi/documents/10623/1107406/Avoimen+tiedon+vaikuttavuus/cd515174-dfeb-4959-a193-246ef2fbefc7?version=1.0>

VM.fi n.d. Haettu 7.4.2017 osoitteesta <http://vm.fi/talouspolitiikka/valtionalouden-kehukset-ja-budjetti>

## KUVALÄHTEET

Euroopan komissio (2015). *Creating Value through Open Data; Study on the Impact of Re-use of Public Data Resources*. Haettu 4.5.2017 osoitteesta [https://www.europeandataportal.eu/sites/default/files/edp\\_creating\\_value\\_through\\_open\\_data\\_0.pdf](https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf)

Gartner (2017). *Magic Quadrant for Business Intelligence and Analytics Platforms*. Haettu 14.4.2017 osoitteesta <https://www.gartner.com/doc/reprints?id=1-3TXXSLV&ct=170221&st=sb>