
DATAVISUALISOINNIT TABLEAU-OHJELMISTOLLA



Ammattikorkeakoulun opinnäytetyö

Tietojenkäsittely

Visamäki, syksy 2016

Atso Arat



VISAMÄKI
Tietojenkäsittely

Tekijä Atso Arat **Vuosi** 2016

Työn nimi Datavisualisoinnit Tableau-ohjelmistolla

TIIVISTELMÄ

Tässä opinnäytetyössä tutkitaan Tableau-ohjelmistoa ja sen toimintoja ja ominaisuuksia verkkopohjaisten datavisualisointien luomiseen. Työn toimeksiantajana toimii Hämeen ammattikorkeakoulun Älykkäät palvelut – tutkimusyksikkö. Tutkimusyksiköllä oli tarve etsiä sopiva työkalu erilaisista datalähteistä tulevan datan visualisointiin niin omaan sisäiseen käyttöön, kuin myös asiakkaille ja yhteistyökumppaneille esitettäväksi.

Työn teoriaosuudessa käydään läpi datavisualisointeja yleisellä tasolla ja käytännönsuudessa tutkitaan mitä ominaisuuksia Tableau-ohjelmistolla on tarjota datavisualisointien luomiseen verkkoon ja onko se ominaisuuksiltaan riittävä toimeksiantajan käyttöön.

Tutkimusmenetelmäksi valittiin kvalitatiivisista menetelmistä haastattelu, jonka avulla selvitettiin toimeksiantajan tarpeita ohjelmiston suhteen. Muutaman kirjan lisäksi, työssä on käytetty pääasiassa lähteinä verkkosivuja. Työn tuloksena voitiin todeta Tableaun vastaavan hyvin toimeksiantajan teknisiä vaatimuksia käyttöön otettavalle ohjelmistolle. Työn pohjalta toimeksiantaja voi ottaa ohjelmiston käyttöönsä ja mikäli se todetaan toimivaksi, voidaan se mahdollisesti ottaa käyttöön muissakin HAMK:n yksiköissä.

Avainsanat Datavisualisointi, Data, Tableau, BI,

Sivut 31 s.



VISAMÄKI
Business information Technology

Author Atso Arat **Year** 2016

Subject of Bachelor's thesis Data Visualizations with Tableau Software

ABSTRACT

This thesis is about Tableau software and its functions and features for creating web based data visualizations. The client for this thesis was the Smart Services research unit from Häme University of Applied Sciences. The research unit had a need to find a proper tool for visualizing data from different kind of sources for their domestic use, and also to show for their clients and associates.

The theory part of this thesis is about data visualizations in general and the practical part is about testing what kind of tool Tableau is for creating web base interactive data visualizations and does it have all the features that the clients wants from this type of software.

Qualitative research methods interview was used for this research to find out the properties that the Smart Services research unit wants the software to have. Besides a few books, almost all my sources for this thesis were web sites. As a result, it can be stated that the Tableau has almost all of the technical features that the client requested for data visualization software. Based on this research, the client can start using the software and if it fits for their workflow and processes, it could be deployed to other units in HAMK as well.

Keywords Data, visualizations, Tableau, BI.

Pages 31 p.



SISÄLLYS

1	JOHDANTO	1
1.1	Tutkimuksen tavoitteet	1
1.2	Toimeksiantajan esittely	2
2	DATAVISUALISOINNIT	3
2.1	Data	4
2.2	Datavisualisoinnin suunnittelu	4
2.2.1	Gestalt laws, hahmolait	5
2.2.2	Hahmolait käytännössä	6
2.3	Visualisoinnit web-ympäristössä	8
2.4	Erilaisia visualisointimalleja	8
2.4.1	Karttapohjaiset visualisoinnit	9
2.4.2	Hierarkiamalliset visualisoinnit	9
2.4.3	Verkostomalliset visualisoinnit	9
3	TABLEAU-OHJELMISTOT	10
4	TUTKIMUSMENETELMÄT JA TUTKIMUKSEN TAVOITE	12
5	VISUALISOINTIEN TOTEUTTAMINEN	13
5.1	Termit	13
5.2	Datalähteet ja datan tuonti	13
5.3	Tietokantaan yhdistäminen	14
5.4	Datatyypit	16
5.5	Sivutyypit	17
5.6	Visualisoinnin rakentaminen - Worksheet	18
5.6.1	Akselittomat mallit	24
5.6.2	Karttapohjaiset mallit	25
5.7	Koostaminen, Dashboard	25
5.8	Esitykset, Storyboard	26
6	VISUALISOINNIN TALLENTAMINEN JA JULKAISU	28
6.1	Tableau server	28
6.2	Server API	30
6.3	Tableau Online	30
7	TUTKIMUKSEN TULOKSET	31
	LÄHTEET	32

1 JOHDANTO

Dataa on kerätty ja sitä on esitetty visuaalisessa muodossa jo pitkälti ennen ajanlaskumme alkua. Tähän aikaan dataa kuvattiin piirtämällä luolien seiniiin tai savilaattoihin. Kerätty data saattoi kertoa eri kuukausien metsästysonnosta tai sääilmiöiden ajankohdista vuosien mittaan. Luolan seinämien sijaan tänä päivänä dataa kerätään digitaaliseen muotoon erilaisiin tietokantoihin ja sitä kerätään jatkuvasti myös enemmän ja enemmän.

Nykypäivänä kerättävästä datasta on loputon määrä esimerkkejä. Kaupat keräävät tietoa asiakkaidensa ostokäyttäytymisestä bonuskorttijärjestelmiensä avulla asiakkaalle tarkkaan kohdennettua mainontaa varten, musiikkipalvelut tutkivat kuuntelijoidensa musiikkimakua, jotta ne osaavat tarjota kuuntelijalle mahdollisesti hänelle sopivaa musiikkia. Myös erilaiset puettavat älylaitteet kuten aktiivisuusrannekkeet ja urheilukellot seuraavat ja tallentavat dataa, jolloin niiden käyttäjä voi tarkastella ja analysoida omia suorituksiaan hyvinkin tarkasti.

Kun on aika tarkastella kerättyä dataa, pelkkiä numeroita selaillessa on hyvin vaikea hahmottaa dataa kokonaisuutena, mutta kun avuksi otetaan jokin visuaalista, jostain helposti ja nopeasti hahmotettavaa ja ymmärrettävää, on tilanne toinen. Kerätystä datasta pystytään muodostamaan selkeä kuva, jolloin erilaisten kuvioiden kuten toistumien löytäminen helpottuu huomattavasti. Ihmiset ovatkin jo aikoja sitten ymmärtäneet, että dataa on huomattavasti helpompaa tulkita, jos se on esitetty visuaalisesti pelkkien numeroiden tai kirjaimien sijaan. Tällöin voidaan puhua datavisualisoinista.

1.1 Tutkimuksen tavoitteet

Tämän opinnäytetyön tavoitteena on tutkia digitaalisessa muodossa olevan datan visualisoinnista käyttäen siihen tarkoitettua datavisualisointiohjelmistoa. Tarkoituksena on tutkia, mitä mahdollisuuksia verkon kautta jaettavat interaktiiviset datavisualisointiesitykset tarjoavat verrattuna perinteisiin kuvamuotoisiin staattisiin kuvaajiin.

Työn teoriaosuudessa käsitellään yleisesti datavisualisointeja: mitä sääntöjä ja käytäntöjä liittyy datan visualisointiin ja erilaisten kuvaajien ja kaavioiden luomiseen, jotta ne eivät vääristele dataa ja katsoja pystyy tulkitsemaan niitä sujuvasti ja oikein. Teoriaosuudessa perehdytään myös siihen, kuinka ihminen tulkitsee visuaalisia elementtejä ja osaa muodostaa visualisoidusta datasta johtopäätöksiä. Työn teoriaosuus pyrkiikin vastaamaan ensimmäiseen tutkimuskysymykseen: Miten rakentaa datavisualisointiesitys oikeaoppisesti?

Käytännön osuudessa tutustutaan Tableau-ohjelmiston ominaisuuksiin ja toimintoihin ja luodaan esimerkkidataa käyttäen erilaisia datavisualisointiesityksiä. Tämän osuuden on tarkoitus vastata toiseen tutkimuskysymykseen: Millainen ohjelmisto on Tableau ja millaisia toiminnallisuuksia se tarjoaa selainpohjaisten datavisualisointien luomiseen.

Toimeksiantajalla on tarve luoda erimallisia datavisualisointeja, joita olisi helppo jakaa edelleen verkon välityksellä. Perinteisten kuvamuodossa jaettujen visualisointien sijaan Tableaulla luodut interaktiiviset visualisointimallit antaisivat mahdollisuuden jakaa informaatiota tehokkaammin asiakkaiden ja yhteistyökumppanien kesken.

1.2 Toimeksiantajan esittely

Opinnäytetyön toimeksiantaja on Hämeen ammattikorkeakoulun älykkäät palvelut -tutkimusyksikkö. Yksikkö sijaitsee Hämeen ammattikorkeakoulun Visamäen toimipisteessä ja se toimii alueellisten yritysten sekä organisaatioiden kanssa yhdessä kehittäen uusia ratkaisuja digitaalisuuden kuin myös palveluliiketoiminnan puolella. (Hämeen ammattikorkeakoulu. n.d-a.)

Hämeen ammattikorkeakoulun eli HAMK:n toiminta on jaettu viiteen yksikköön, joiden tehtävänä on järjestää opetusta, tehdä tutkimusta, ylläpitää kansainvälistä toimintaa sekä toimia yhteistyössä alueellisten yritysten kanssa (Hämeen ammattikorkeakoulu. n.d-b.)

HAMKiin kuuluu seitsemän kampusta, joissa yhteensä noin 30 koulutusohjelmaa. Se tarjoaa mahdollisuuden suorittaa AMK-tutkinnon kuin myös ylemmän AMK-tutkinnon päivätutkintona tai aikuisopiskeluna. (Hämeen ammattikorkeakoulu. n.d-c.)

2 DATAVISUALISOINNIT

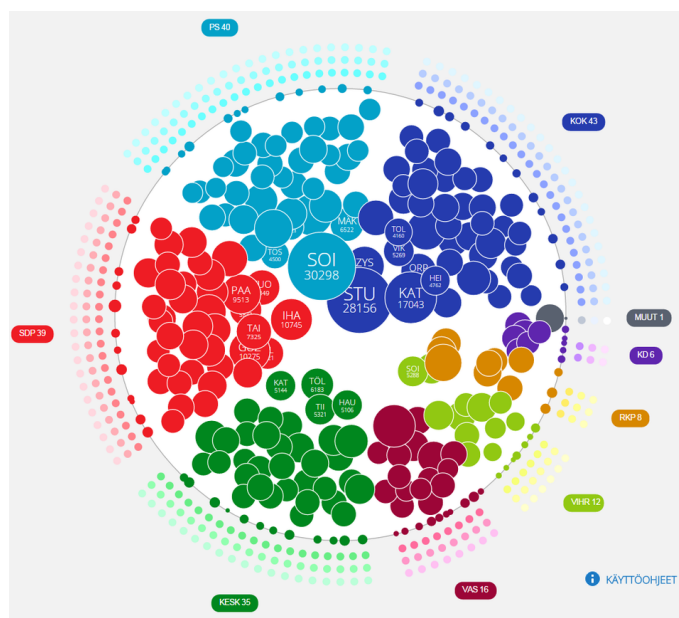
Dataa kerätään nykyisin todella paljon ja monista eri lähteistä, mutta siitä huolimatta se on merkityksellistä vain, kun siitä voidaan saada irti tietoa. Dataa pitää käsitellä ja jalostaa, jotta se saadaan käyttökelpoiseen ja ymmärrettävään muotoon. Raaka numeerinen data on jo sinällään melko vaikeaa tulkittavaa, mutta siitä trendien tai toistuvien kuvioiden löytäminen melkein mahdotonta. (Murray 2013, 1.)

Kun apuna käytetään visuaalisia elementtejä, voi kuka tahansa helposti tulkita näkemäänsä dataa. Datavisualisointien tarkoitus on esittää moniulotteinen ja laaja datamäärä helposti ja nopeasti omaksuttavassa muodossa. Niiden avulla dataa voidaan esittää havainnollisesti, laadukkaasti sekä dynaamisesti ja vuorovaikutteisesti hyödyntäen erilaisia animaatioita ja graafisia elementtejä. (Kettunen 2015; Murray 2013, 1.)

Datavisualisointi käytännössä tarkoittaa datan sitomista visuaalisiin elementteihin. Numeerisen datan tietty arvo voi tarkoittaa esimerkiksi perinteisessä palkkikaaviossa palkan pituutta tai vaikka ympyröistä koostuvassa visualisoinnissa pallon halkaisijaa tai sen kehän paksuutta. (Murray 2013, 1-2.)

Perinteisen malliset datavisualisoinnit kuten palkki -ja pylväsdiagrammit riittävät monesti yksinkertaiselle ja pienellä datamäärälle, mutta kun dataa on paljon ja halutaan tarkastella lähemmin muuttujien välisiä suhteita, on olemassa myös moniulotteisempia malleja. (Tampereen yliopisto. n.d-a.)

Esimerkkinä voidaan mainita YLE, joka on hyödyntänyt erikoisempia datavisualisointimalleja useasti. Esimerkiksi vuonna 2012 presidentin vaaleissa luomalla ehdokkaiden Twitter-seuraajista kaavion, sekä kaksissa edellisissä eduskuntavaaleissa ehdokkaiden äänimäärät esittävän ”vaalipallon”. (YLE. 2012; 2015.)



Kuva 1. Ylen vaalipallo vuoden 2015 eduskuntavaaleista. (YLE 2015)

2.1 Data

Datalla tarkoitetaan yleensä jonkinlaista raakaa dataa joka ei itsessään ole tulkittavassa muodossa, mutta sitä on kuitenkin mahdollista jalostaa ymmärrettävään muotoon. Data voi olla mitä tahansa, merkkejä, symboleja tai vaikka numeroita, mutta sinällään tarkasteltuna sen pohjalta ei pystytä vielä muodostamaan mitään ymmärrettävää. Kun data on analysoitu ja tulkittu saadaan se ihmiselle ymmärrettävään muotoon, tällöin voidaan sanoa syntyneen informaatiota. Informaatio on tietoa, jolla on merkitys sekä sisältö. (Haasio & Vakkari n.d-a; Jokela 2011.)

Kun puhutaan digitaalisesta datasta, pitää sen tietysti olla tallennettu jontekin. Yksi tyypillisimmistä vaihtoehdoista on erilaiset laskentataulukot. Ehkä tyypillisin on Excel-taulukko tai jokin muu vastaava taulukkoformaatti kuten CSV tai TSV. Viimeksi mainitut taulukkoformaattit eroavat tavallisesta Excel-taulukosta lähinnä vain eri tyyppisellä dataelementtien erottelulla. (Creativyst n.d-a.)

Toinen vaihtoehto datan säilyttämiseen on tietokanta. Tyypillisiin tietokantatyyppeihin on relaatiotietokanta. Se muodostuu useista tauluista, jotka rakentuvat sarakkeista ja riveistä. Relaatiotietokannassa yksittäinen tieto tallennetaan vain yhteen paikkaan ja taulujen välille luodaan relaatioita eli yhteyksiä, jotka auttavat minimoimaan tietokannan tekemän työn. Tällöin muutoksia tarvitsee tehdä vain yhteen paikkaan kerrallaan johtuen taulujen yhteyksistä. (Sarja 2006.)

2.2 Datavisualisoinnin suunnittelu

Datavisualisoinnit hyödyntävät aivojen visuaalista havainnointikykyä. Ihminen pystyy käsittelemään näkemänsä hyvin tehokkaasti ja voimme välittömästi ymmärtää visuaalisesti esitettyä dataa. Jos tarkastelussa taas on pelkkiä numeroita, aivot tukeutuvat kognitiiviseen puoleen ja tällöin kokonaisuuden hahmottaminen vie huomattavasti enemmän aikaa. (Creativebloq 2013.)

Vaikka kyse onkin datan muuttamisesta visuaaliseen muotoon, ei sen pääasiallinen tehtävä ei kuitenkaan ole ulkonäkö. Datavisualisoinnin tehtävä ei ole näyttää kauniilta, vaan välittää tieto tehokkaasti ja vääristelemättä katsojalleen. Tällä tarkoitetaan sitä, että jokainen graafisen elementin, värin tai muun yksityiskohdan tulisi olla merkityksellinen esitettävän datan kannalta. (informaatiomuotoilu.fi 2012.)

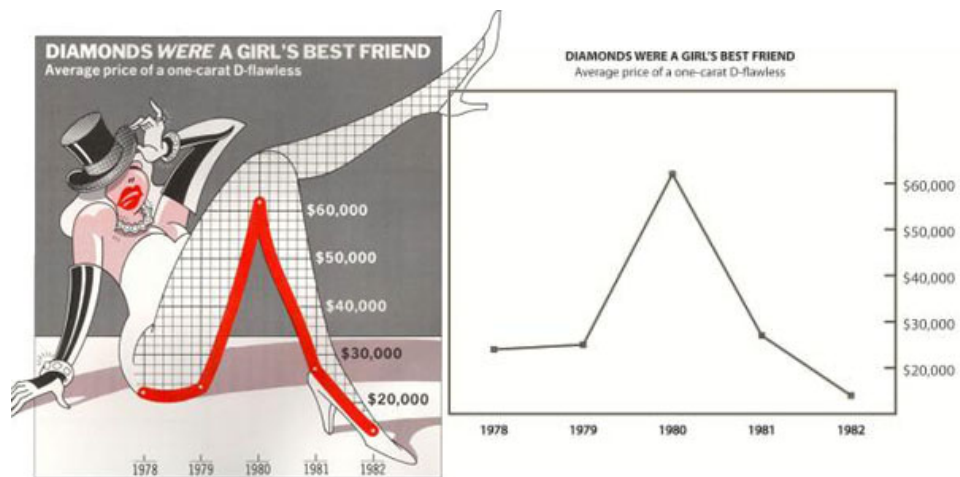
Moni seikka vaikuttaa datavisualisoinnin ulkoasun suunnitteluun. Näistä yksi tärkeä on esityksen kohderyhmä. Mikäli valitaan jokin erikoinen visualisointimalli jota kohdeyleisö ei välttämättä osaa tulkita riittävän hyvin ja nopeasti, voi sen välittämä viesti mennä katsojalta täysin ohi. Toisaalta taas yleisön liiallinen aliarviointikaan ei välttämättä ole hyväksi. (Kuusela 2000, 200) Tämän seikan takia, voi olla hyvä miettiä dataesitystä luodessa, onko se suunnattu monelle eri tasoisille kohdeyleisöille vai vain ainoastaan yhteen tiettyyn esitykseen yhdelle tietylle yleisölle.

Ei ole täysin itsestään selvää mitkä ominaisuudet tekevät hyvän tai huonon datavisualisoinnin. Jossain tapauksissa joku toimii ja hyvin ja toisessa ei ollenkaan. Kuitenkin on tiettyjä ominaisuuksia mitä pidetään yleisesti määritelmänä hyvälle ja oikeaoppiselle datavisualisoinnille. (Kuusela 2000, 24; Interaction design foundation n-d.a.)

- Ilmaistaan selkeästi osien suhteet kokonaisuuteen.
- Osien vertailu keskenään on helppoa.
- Määrät on esitetty selkeästi.
- Määrien vertaileminen on helppoa.
- Kertoo selkeästi mitä kuvattu data on ja mikä on sen tarkoitus.

(Interaction design foundation n-d.a; Datalabs n-d.a.)

Toisaalta taas, mikäli dataa ei itsessään vääristellä, erittäin graafiset visualisoinnit saattavat jäädä ihmisten mieleen huomattavasti paremmin kuin pelkistetyt. Saskatchewanin yliopiston tekemän tutkimuksen mukaan, niin sanotusta ”roskagrafiikasta” tieto välittyi aivan yhtä hyvin, lukijat muistivat kaavion sisällön huomattavasti paremmin kuin minimalistisesta verrokistaan ja yleisesti pitivät niitä mielekkäämpinä katseltavina. (University of Saskatchewan 2010)



Kuva 2. ”Roskagrafiikka”. Graafinen ja minimalistinen kuvaaja. (University of Saskatchewan 2010)

2.2.1 Gestalt laws, hahmolait

Ihminen tulkitsee luonnostaan tiettyjä asioita tietyllä tavalla. Esimerkiksi numerot eivät luonnostaan kerro kenellekään mitään, sillä ne ovat opeteltuja symboleja. Kun ihminen katsoo lyhyttä ja pitkää viivaa, hän ymmärtää heti mistä on kysymys. Pidempi viiva kertoo suuremmasta arvosta kuin lyhyt. Muotoja taas ei voida käyttää kuvaamaan määrällisiä arvoja, mutta niillä voidaan kuvata erottelu, ryhmittelyä ja yhteyksiä. (Creativebloq 2013.) Tähän pätee ns. Gestaltin lait eli Hahmolait.

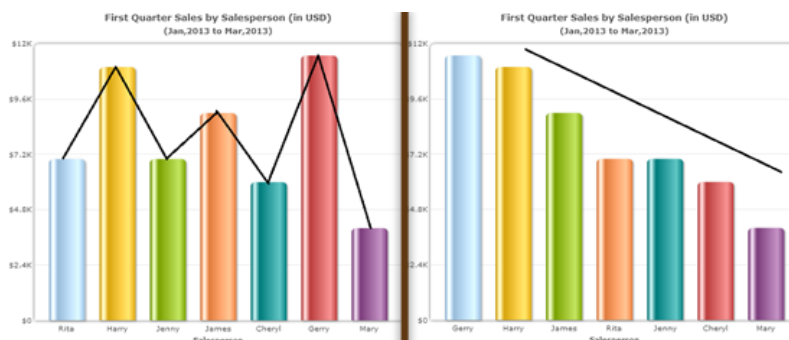
Gestalt-teoria syntyi vuonna 1890 Saksassa. Se kuvaa sitä, miten näköaistin kautta havaitut yksittäiset objektit voidaan havaita osana laajempaa kokonaisuutta. Tähän liittyvät ns. Gestaltin hahmolait, jotka kuvaavat millä kaikilla tavoin tämä vaikutus voidaan saada aikaan. Lakeja on melko paljon, mutta seuraavassa esiteltynä keskeisimmät käsitellyyn aiheeseen liittyen. (Jyväskylän yliopisto 2004.)

- Samankaltaisuuden (Law of similarity) lain mukaan. Kuviot jotka ovat väreiltään tai muodoiltaan yhdenkaltaiset voidaan mieltää kuuluvan samaan ryhmään.
- Läheisyyden (Law of proximity) lain mukaan toisiaan lähellä olevat kappaleet voidaan mieltää kuuluvan samaan ryhmään.
- Yhteenliittymisen (Law of unity) lain mukaan kuviot jotka ovat jollain tavalla yhdistetty toisiinsa kuuluvat samaan ryhmään.
- Sulkeutuvuuden (Law of closure) lain mukaan kappaleita ympäröivä täysin tai melkein suljetun viivan muodostama muoto muodostaa ryhmän.
- Jatkuvuuden (Law of continuity) lain mukaan monesta erillisestä kappaleesta rakentuva viiva tai kuvio mielletään yhtenäiseksi.
- Keskipisteen (Law of focal point) lain mukaan yksittäinen selkeästi ympäristöstään erottuva elementti keskittää katsojan huomion tähän kohtaan. (Jyväskylän yliopisto 2004; Padmanabhan 2013.)

2.2.2 Hahmolait käytännössä

Kun näitä sääntöjä sovelletaan käytännössä datavisualisointeihin, saattavat ne tuntua monelle itsestäänselvyyksiltä, mutta kaikesta huolimatta ne vaikuttavat huomattavan paljon katselukokemukseen ja esitettävän datan ymmärtämiseen.

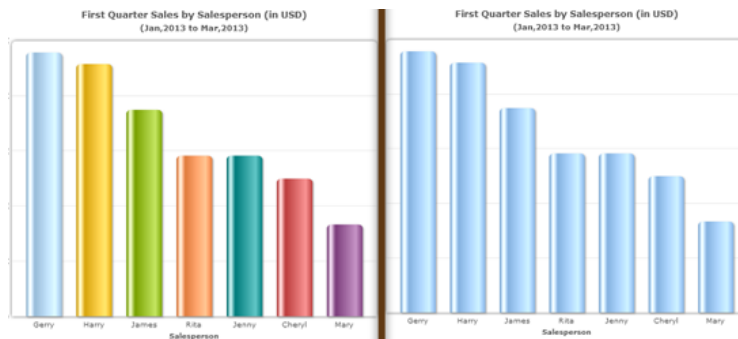
Monta yksittäistä muotoa voidaan havaita yhtenä tai ryhmänä, mikäli ne ovat järjestelty oikein. Kuvan 3 esimerkistä voi selkeästi huomata, että oikeinpuoleinen kuvaaja on huomattavasti helpolukuisempi, kun palkit ovat järjestelty pituusjärjestykseen. (Padmanabhan 2013.)



Kuva 3. Jatkuvuus pylväskaaviossa. (Padmanabhan 2013)

Samankaltaisuuden lain mukaan objektit, jotka ovat samankaltaisia väreiltään, muodoiltaan, kooltaan tai asennoltaan, mielletään ryhmäksi.

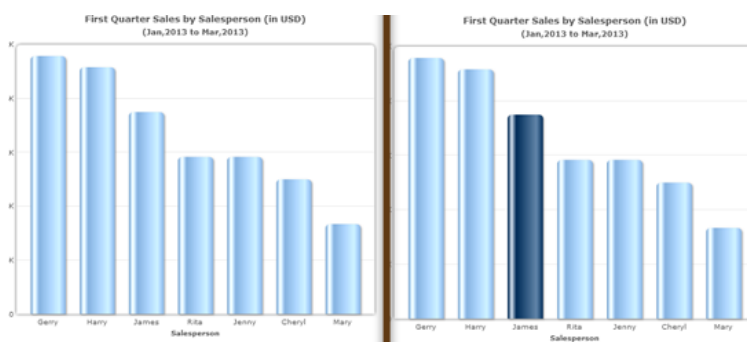
Kuvassa 4 näkyvässä vasemman puoleisessa kaaviossa värit eivät kuvaa mitään dataa, vaan ne ovat pelkästään kosmeettisista syistä käytössä. Kun poistetaan turha häiritsevä ja merkityksetön elementti kaaviosta, voidaan varmistua, että se ei vaikuta harhaanjohtavasti kaavion tulkinnassa. (Padmanabhan 2013.)



Kuva 4. Samankaltaisuus pylväskaaviossa. (Padmanabhan 2013)

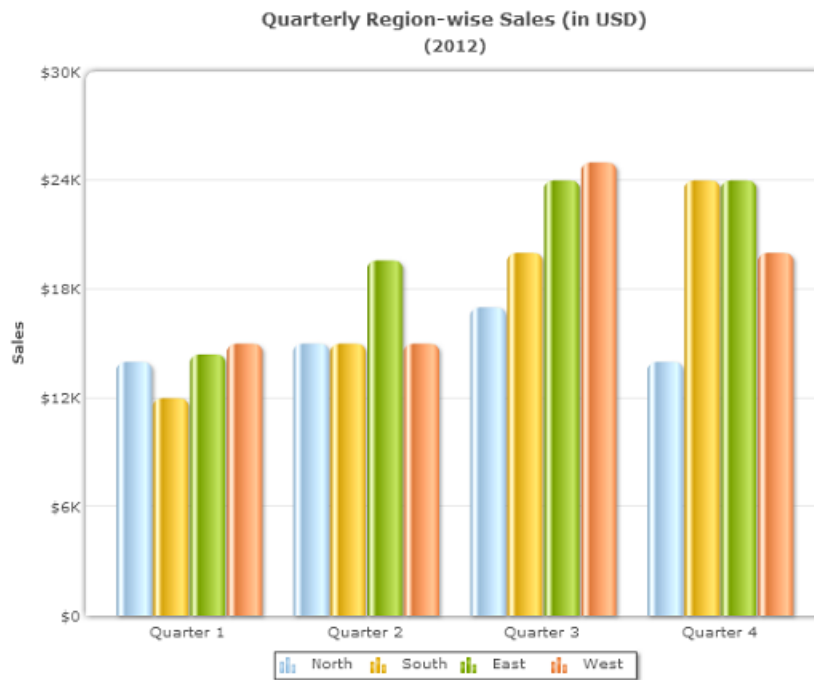
Keskipisteen lain mukaan yksittäinen joukosta erottuva elementti vetää katsojan huomion puoleensa. Tällä tavoin voidaan antaa katsojalla aloituskohta josta lähteä tutkimaan dataa pidemmälle. Tämä on hyvä keino erottaa joukosta jokin merkitsevä elementti, joka on keskeinen asia kuvaajasaa. (Padmanabhan 2013.)

Kuvan 5 esimerkissä haluttu yksityiskohta on nostettu esiin väriä käyttämällä. Tämä toimii hyvin, mikäli data tarkoituksella esitetään jostain tietystä perspektiivistä ja halutaan verrata tätä tiettyä elementtiä muihin.



Kuva 5. Keskipisteen laki tuo esille halutun yksityiskohdan (Padmanabhan 2013)

Läheisyyden lain mukaan objektit jotka ovat sijoitettu lähelle toisiaan nähdään ryhmänä. Näin voidaan helposti vertailla arvoja ryhmän sisällä ja sen jälkeen siirtyä vertaamaan ryhmiä keskenään. Kuvassa 6 on ryhmitteilyä hyödynnetty erottamaan vuosineljännekset toisistaan. (Padmanabhan 2013.)



Kuva 6. Ryhmittely hyödyntäen läheisyyden lakia. (Padmanabhan 2013)

2.3 Visualisoinnit web-ympäristössä

Verkko tarjoaa alustana monia erilaisia ominaisuuksia ja mahdollisuuksia datan esittämiseen, mitä yksinkertainen kuvatiedosto tai tuloste eivät pysty tarjoamaan. Nykyisin se on myös huomattavasti tehokkaampi tapa tiedonjakoon esimerkiksi sosiaalista mediaa hyödyntäen.

Erilaiset web-tekniikat tarjoavat paljon mahdollisuuksia. Dataesitykseen voidaan lisätä animaatiota ja tehosteita ja siitä voidaan tehdä interaktiivinen, jolloin käyttäjä pääsee itse tutkimaan dataa laajemmin ja haluamallaan tavalla. Verrattuna perinteisiin staattisiin visualisointeihin, joissa jokainen eri tilanne vaatii oman erillisen kaaviokuvansa, voidaan verkossa luoda selaimella toimiva yksittäinen visualisointikonaisuus, joka on interaktiivinen ja pitää jo sisällään kaiken tarvittavan datan eri mallien koostamiseen. Tämä mahdollistaa dataa tarkastelevalla monipuolisen selailukokemuksen aineistoon ja mahdollistaa datan tarkastelun juuri halutusta perspektiivistä. (Murray 2013, 3.)

2.4 Erilaisia visualisointimalleja

Dataa on hyvin monenlaista ja niin on myös erilaisia visualisointimalleja. Käytettävään visualisointimalliin vaikuttaa se millaista käytettävä data on ja se mitä sen avulla halutaan kertoa. Yleensä visualisoinneissa ja kaavioissa halutaan esittää jakautumista, trendejä, vertailua, tai sijaintia. (My market research methods 2013.) Tämän työn puitteissa en käsittele kaikkia mahdollisia visualisointimalleja, mutta seuraavassa mainittuna muutamia erityyppisiä visualisointityyppejä.

2.4.1 Karttapohjaiset visualisoinnit

Karttapohjaisia malleja on monia erilaisia ja ne ovatkin ehkä suurin alakategoria visualisointimalleissa. Näiden kaikkien perusajatus on kuitenkin sama: ne kaikki piirtävät jollain tavoin karttapohjan päälle joko maalamalla alueita tai sijoittamalla kohteita kartalle.

Tällä tavoin voidaan ilmaista esimerkiksi jonkin asian tiheyttä tai määrää tietyllä maantieteellisellä alueella. Datan pitää luonnollisesti pitää sisällään jonkinlaista paikkatietoa, yleensä koordinaatteja, jotta tämän tyyppisen mallin luominen on mahdollista. (Datalabs n.d-a; Skau 2012.)

2.4.2 Hierarkiamalliset visualisoinnit

Tämän tyyppin malleilla voidaan kuvata asioiden ryhmittelyä ja järjestystä. (Datalabs n.d-a.) Esimerkiksi organisaation tai yrityksen hallinto voidaan kuvata hyvin tämän tyyppisellä kaaviolla.

2.4.3 Verkostomalliset visualisoinnit

Tämän tyyppiset mallit kuvaavat asioiden yhteyksiä toisiinsa. Niissä voidaan esittää tiettyjen asioiden yhteyksiä kuin myös mikä niitä yhdistää. (Datalabs n.d-a.)

3 TABLEAU-OHJELMISTOT

Tableau on 2003 Kaliforniassa Yhdysvalloissa perustettu, nykyisin pääkonttoriaan Seattlessa Yhdysvalloissa pitävä ohjelmistoyhtiö, jonka tuotteet keskittyvät pääasiassa datavisualisointeihin ja liiketoimintatiedon hallintaan. Kirjoitushetkellä yhtiö työllistää yhteensä noin 2400 ihmistä ja sillä on yli 23 000 yritysasiakasta, joihin lukeutuu mm. Audi, SpaceX, Cisco ja Citrix. (Tableau n-d.a.)

Yhtiön perustivat Chris Stolte, Christian Chabot ja Pat Hanrahan. Alun perin he tutkivat datavisualisointitekniikoita Stanfordin yliopistossa Yhdysvaltain puolustusministeriölle. He loivat datavisualisointiohjelmiston tutkimuksensa pohjalta ja kaupallistivat toimintansa. (Tableau n.d-b.)

Opinnäytetyön kirjoitushetkellä Tableau ohjelmistoperheeseen kuuluu yhteensä viisi eri tuotetta, joista tässä opinnäytetyössä käsitellään lähinnä Tableau desktop -työpöytäsovellusta ja sen käyttöä ja ominaisuuksia. Työssä käytettiin ohjelmiston 9.2-versiota, joka oli työn aloitus hetkellä uusien saatavilla oleva.

Tableau desktop on työpöytäsovellus ja tuoteperheen keskeisin sovellus jonka avulla varsinaiset visualisointikonaisuudet rakennetaan. Sitä tarjotaan kolmena eri versiona: Public, Personal ja Professional. Näistä Public on yksityiskäyttöön tarkoitettu täysin ilmainen, mutta jokseenkin rajoitettu versio. Näistä vain professional-versiota on mahdollista käyttää yhdessä Server- ja Online-ohjelmistojen kanssa. Ohjelmistot ovat saatavilla niin Windows kuin Mac OS ympäristöönkin. (Tableau n.d-c.)

Tableau Server on palvelinohjelmisto, johon työpöytäsovelluksella luodut visualisoinnit voidaan sijoittaa ja sitä kautta jakaa esiteltäviksi tai muokattaviksi. Palvelinohjelmisto on saatavilla vain Windows alustalle ja sitä käytetään verkkoselaimen kautta. Android ja IOS alustoille on myös olemassa omat sovellukset, jotka helpottavat palvelimen käyttö mobiililaitteilla. (Tableau n.d-d.)

Tableau Online on Tableau Server tarjottuna SaaS-mallisena palveluna ja se pitää sisällään melkein kaikki Server-ohjelmiston ominaisuudet. Pieniä eroja löytyy, esimerkiksi oman sivustokokonaisuuden brändääminen omilla logoilla ja grafiikoilla ei ole mahdollista, sekä sisällön tarkasteleminen vaatii aina kirjautumisen palveluun. Vieraskirjautuminen ilman tunnuksia ei siis onnistu. (Tableau n.d-e.)

Tableau reader on ilmainen työpöytäsovellus, jolla voidaan avata ja tarkastella Tableau omaan projektitiedostomuotoon tallennettuja visualisointeja. Sovellus osaa vain lukea tiedostoja, joten minkäänlaisia muokkaamismahdollisuuksia ei käyttäjällä tässä tapauksessa ole. (Tableau n.d-f.)

Ohjelmistojen hinnoittelu oli kirjoitushetkellä seuraavanlainen: Tableau desktop ohjelmistolle on kaksi erilaista lisenssiä, personal ja professional. Personal on huomattavasti rajoitetumpi lisenssi ja se myydään käyttäjäkohtaisesti. Personal-lisenssin hinta on 999\$. Professional on ns. täysver-

sio ja sen hinta on 1999\$, professional ei ole rajoitettu yhteen käyttäjään. (Tableau n.d-g.)

Tableau server-ohjelmiston hinta on 10 000\$ ja se sisältää 10 käyttäjää. Tarkempia hintatietoja ei Tableau sivuillaan kerro. Tablea Online-palvelun hinta on 500\$ per käyttäjä/per vuosi. Tableaulla on myös olemassa oma hinnasto opiskelijoille ja oppilaitos käyttöön. (Tableau n.d-g.)

4 TUTKIMUSMENETELMÄT JA TUTKIMUKSEN TAVOITE

Tutkimuksen tavoitteena on tutkia Tableau-datavisualisointiohjelmistoa ja sen toimintoja sekä ominaisuuksia. Tarkoituksena tutkia ohjelmiston sopivuutta Älykkäät palvelut –tutkimusyksikön käyttöön.

Käytössä tässä opinnäytetyössä olivat kvalitatiivisista tutkimusmenetelmistä haastattelu, jonka avulla selvitettiin toimeksiantajan tarpeita ja vaatimuksia ohjelmiston ja sen toiminnan suhteen sekä käsiteltiin tällä hetkellä käytössä olevia ratkaisuja. Tämä opinnäytetyö toimii samalla esittelynä Tableusta ja sen mahdollisuuksista, jonka perusteella toimeksiantaja voi harkita ohjelman käyttöönottoa.

Haastateltavina olivat tämän työn toimeksiantajan eli Älykkäät palvelut – tutkimusyksikön tutkimuskoordinaattori Joni Kukkamäki sekä tietojenkäsittelyn opettaja Turo Nylund, joka myös työskentelee tutkimusyksikössä.

Haastattelun avulla selvitettiin mitä tekniikoita toimeksiantajalla on käytössään tällä hetkellä liittyen datan keräämiseen ja datavisualisointien luomiseen ja mitkä ovat ne tekniikat jotka Tableausta tulisi löytyä, jotta se sopisi nykyisiin työskentelymenetelmiin ja malleihin.

Ohjelmiston mahdollinen rooli olisi toimia niin organisaation sisäisessä kuin ulkoisessakin käytössä. Sen tulisi olla työkalu, jonka avulla voitaisiin nopeasti jakaa dataa erilaisista tutkimustuloksista sisäisesti ja tähän Tableaun tarjoamat verkkopohjaiset ja moniulotteiset visualisointiesitykset voisivat sopia hyvin.

Organisaation ulkopuolelle jaettava data ainakin alkuun koostuisi todennäköisesti vain perinteisistä staattisista esityksistä. Mikäli ohjelmistolla voidaan kasata hyvännäköisiä monen graafin kokonaisuuksia, jotka voitaisiin tallentaa ihan vaikka kuvatiedostona kuten MS Excelissäkin, olisi se toimiva.

Haastattelussa Nylund listasi tärkeimmiksi asioiksi tuen erilaisille tietolähteille nimeten erityisesti MySQL sekä MongoDB tärkeiksi tekniikoiksi. Myös tuen Excel ja CSV tiedostoille hän listasi hyvin tärkeäksi. (Kukkamäki & Nylund, haastattelu 18.1.2016)

Nylund ja Kukkamäki vielä lisäsivät, että mikäli tulevaisuuden projekteissa käsitellään big dataa, voi Hadoopin tietokantaratkaisut sekä MS Azure ja Amazonin pilvipalvelut tulla ajankohtaisiksi, joten Tableaun olisi hyvä tukea myös näitä alustoja. (Kukkamäki & Nylund, haastattelu 18.1.2016)

Tiedustelin myös, että minkälaisia projekteja tai testauksia heillä tälle hetkellä on menossa ja missä dataa kerätään ja millä laitteilla. Nylund kertoi, että Thingsee-anturilaitte on sellainen mitä on viime aikoina testattu paljon ja sitä on hankittu useita ympäri HAMKia erilaisiin projekteihin. (Kukkamäki & Nylund, haastattelu 18.1.2016)

5 VISUALISOINTIEN TOTEUTTAMINEN

Työssä käytetään muutamaa eri esimerkkidataa, sekä Tableaun omia esimerkkidatoja. Ensimmäinen on Espoon kaupungin julkaisemaa dataa kaupungin ostoista vuosilta 2012-2014. Data on valmiiksi järjesteltyä taulukodataa Excel-muodossa. Sen sarakkeet ovat: Kaupunki, tuloyksikkö, kustannuspaikka, kustannuspaikan nimi, kirjauspv, laskunnumero, tiliryhmä, tilinumero, tilin nimi, toimittajan nimi, Y-tunnus, euroa, ja tilausnumero. Rivit taas koostuvat yksittäisistä ostotapahtumista. Aineisto on melko kookas ja sitä on rajattu moniin esimerkkikaavioihin.

Toisena esimerkkidatana on toimeksiantajalla käytössä Thingsee-laitteesta kerätty data. Laite on toimeksiantajalla tällä hetkellä vasta alustavassa testauskäytössä, joten data ei ole kerätty mistään merkityksellisestä kohteesta. Thingsee on laite, jossa on monia erilaista dataa kerääviä antureita kuten lämpö – ja valoisuusanturit sekä kiihtyvyyssanturi. Datassa on kolme eri elementtiä: sid (anturin tyyppi), val (anturin arvo) ja ts (aikaleima).

Thingsee ei osaa suoraan lähettää kerättyä dataa mihinkään tietokantaan vaan siihen pitää tehdä hakuja rajapinnan kautta ja tätä kautta siirtää data haluttuun lopulliseen sijoituspaikkaan tätä kautta.

5.1 Termit

Koska Tableau ohjelmistoista ei ole Suomenkielisiä versioita saatavilla, on tässä työssä käytetyt suomenkieliset termit valikoista ja toiminnoista epävirallisia. Käytetyt termit eivät yritä viitata mahdollisiin suomenkielisiin virallisiin talous-tai tilastotieteen termeihin millään tavalla, vaan ne on yritetty kääntää Englannista Suomeksi sanakirjaa käyttämällä. Kaikkia termejä ei kuitenkaan ole käännetty, vaan selvyuden vuoksi ne esitetään alkupe-raisillä nimillään englanniksi.

5.2 Datalähteet ja datan tuonti

Kun aloitetaan visuaalisen ulkoasun luomista datalle Tableaussa, ensin kerrotaan ohjelmalle missä käsiteltävä data sijaitsee. Ohjelmistoon voidaan tuoda dataa kahdella tavalla. Joko Suoraan tiedostosta tai yhdistämällä ohjelma tietokantaan. Ohjelma osaa tulkita monia tavallisimpia tiedostoformaatteja kuten TSV, CSV ja TXT tiedostoja kuin myös Excel tiedostoja sekä Access-tietokantoja.

Tableau tukee melko laajaa valikoimaa erilaisia tietokantayhteyksiä, jotka ovat esitettynä taulukossa 1. Ohjelmassa on myös ”other database” toiminto, jolla voidaan kokeilla yhdistää tietokantaan jota ei ole suoraan tuettu.

Nylund mainitsi haastattelussa MySQL sekä MongoDB olevan heidän kannaltaan keskeisiä tekniikoita datankeräämisessä ja Tableau tukeekin näitä molempia tekniikoita. MongoDB yhdistetään Tableauseen käyttäen PostgreSQL yhdistäjää. Tableau osaa myös yhdistää Hive-tyyppisiin

tietokantoihin kuten esim. Hadoop Hive-kantaa. (Kukkamäki & Nylund, haastattelu 18.1.2016)

Taulukko 1. Tuetut tietokantayhteydet datan tuontiin.

Tableau Server	Action Vector	Amazon Aurora
Amazon EMR	Amazon Redshift	Aster Database
Cloudera Hadoop	DataStax Enterprise	EXASolutions
Firebird	Google Analytics	Google BigQuery
Google Cloud SQL	Hortonworks Hadoop Hive	HP Vertica
IBM BigInsights	IBM DB2	IBM Netezza
MapR Hadoop Hive	MarkLogic	Microsoft Analysis Services
Microsoft PowerPivot	Microsoft SQL Server	MySQL
OData	Oracle	Oracle Essbase
ParAccel	Pivotal Greenplum Database	PostgreSQL
Progress OpenEdge	Salesforce	SAP HANA
SAP NetWeaver Business Warehouse	SAP Sybase ASE	SAP Sybase IQ
Spark IQ	Spark SQL	Splunk
Teradata	Teradata OLAP Connector	

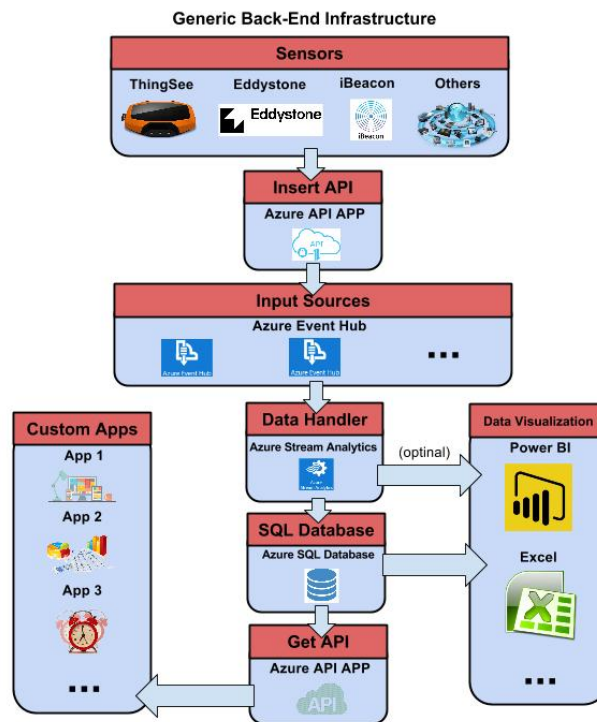
5.3 Tietokantaan yhdistäminen

Thingsee -laitteen keräämä data, jota käytetään yhtenä esimerkkidatana, sijaitsee Microsoftin Azure -pilvipalvelussa Microsoft SQL tietokannassa. Tietokannan data on kerätty Thingsee-laitteesta, jolla on testaus mielessä kerätty erityyppistä anturidataa kuten lämpötilaa, valoisuutta ja kiihtyvyyttä.

Thingsee on itsenäinen akkuvirralla toimiva monianturinen laite, joka voidaan määrittää keräämään erilaista tietoa ympäristöstään kuten esim. lämpötilaa, valoisuutta ja ilmankosteutta. Laite voidaan määrittää lähettämään keräämänsä data eteenpäin joko Wifi verkossa tai mobiilidatayhteyttä.

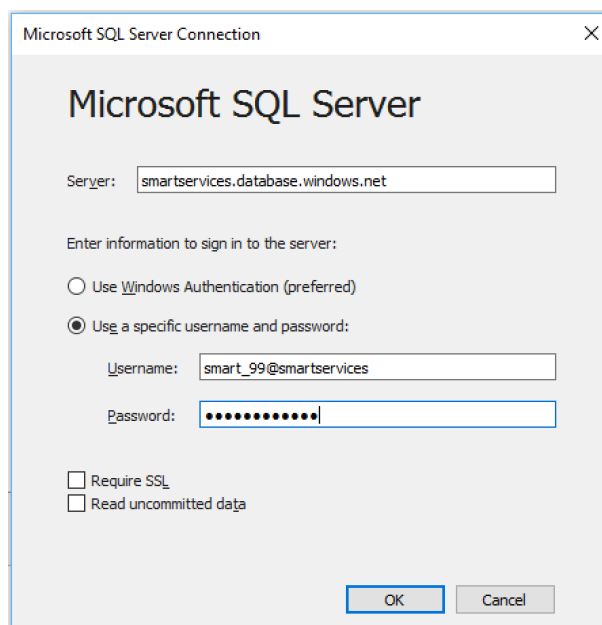
Kuvasta 7 käy ilmi mitä kautta data on tuotu tietokantaan. Kuvassa näkyy myös toimeksiantajan tämän hetkiset ratkaisut datavisualisointien luomiseen, MS PowerBI sekä Excel. Tableaun on mahdollisesti tarkoitus tulla

olemassa olevien ratkaisujen lisäksi tai jopa korvata ne datavisualisointien luomisessa ja jakamisessa.



Kuva 7. Datankeräys-malli. (Älykkäät palvelut –tutkimusyksikkö 2016)

Kun lähdetään luomaan uutta visualisointimallia, ensimmäisenä valitaan käytettävä data. Uusi datalähde lisätään ”Add new data source” painikkeesta. Seuraavaksi valitaan käytetyn tietolähteen tyyppi, joka tässä tapauksessa on Microsoft SQL-tietokanta ajettuna Azuressa. Annetaan palvelun osoite sekä tunnistautumistiedot kuten kuvassa 8 näkyy.



Kuva 8. Yhteyden muodostaminen tietokantaan.

Kun yhteys on onnistuneesti luotu, voidaan valita tietokannasta haluttu taulu, tai taulut joita halutaan tarkastella. Data source ikkuna mahdollistaa taulujen liittämisen yhteen, mikäli tarvittava data on sijoitettu moniin eri tauluihin.

Tässä kohtaa voidaan myös määrittää, miten Tableau hakee dataa käyttöönsä. Vaihtoehtoina on live-tila, jossa uutta dataa haetaan määritellyltä palvelimelta jatkuvasti.







Toisena mahdollisuutena on extract –tila, jossa datasta haetaan aina määräjain uusi kopio Tableaun palvelimelle, jossa lopullinen visualisointikin sijaitsee. Tämä ominaisuus liittyy enemmän visualisoinnin siirtämiseen palvelimelle ja sieltä käyttämiseen, joten tätä ominaisuutta on avattu enemmän luvussa 6.

5.4 Datatyypit

Tableau jakaa datan kahteen eri päätyyppiin: dimensions ja measures, kvalitatiivisiin ja kvantitatiivisiin tyyppeihin. Tableau tulkitsee tuodun datan automaattisesti ja oletuksena kaikki numeroista koostuvat kentät asetetaan kvalitatiivisiksi eli laskennallisiksi päivämääriä ja kellonaikoja lukuun ottamatta, nämä Tableau osaa useasti tunnistaa oikein.

Kaikissa tilanteissa ohjelma ei välttämättä osaa tulkita kaikkia kenttiä oikein ja käyttäjän täytyy korjata niille oikea tyyppi. Esimerkiksi postinumeroilla tai puhelinnumeroilla ei ole mitään laskennallista arvoa, joten ne voitaisiin hyvin muuttaa teksti-tyypiksi. Datan tyyppiä voidaan muuttaa heti tuonti-ikkunassa tai myöhemmin visualisointi-ikkunassa.

Taulukossa 2 on esitetty kaikki kuusi datatyyppiä, joita ohjelmasta löytyy. Sininen tarkoittaa ei-laskennallista ja vihreä laskennallista tyyppiä.

kuvake	Kuvaus
	Teksti
	Päivämäärä
	Päivämäärä ja aika
	Numero
	Totuusarvomuuttuja
	Sijaintitieto

Taulukko 2. Datatyypit

Tableau ei luonnollisesti voi tunnistaa kaikkea dataa suoraan sellaisenaan ja monesti voikin olla tarpeellista luoda ”laskettu kenttä” (calculated field) jossa luodaan olemassa olevasta dataelementistä uusi ajamalla se laskutoimituksen tai funktion lävitse.

Esimerkki tapauksessa Thingsee-laite merkitsee datan tapahtuma ajankohdan Linux EPOCH aikakoodilla. Tableau ei tunnista tätä ajaksi, vaan tulkitsee kentän pelkkänä numeroina. Jotta aikakoodi saadaan ymmärrettävään muotoon oikeaksi päivämääräksi ja ajaksi, täytyy sille luoda laskennallinen datakenttä.

Kuvassa 10 on käytetty dateadd-funktiota, jolla aikakoodi on muutettu ymmärrettävään muotoon. Koodissa oranssina merkattu "ts" on alkuperäisen aikakoodikentän nimi.



Kuva 9. laskennallinen datakenttä "aika" luodaan.

Uusi luotu kenttä voidaan nyt määrittää joko pelkäksi päivämääräksi tai päivämäärä ja aika-kentäksi. Tällöin ohjelma osaa käsitellä sitä oikein.

5.5 Sivutyypit

Kun ollaan yhdistetty käytettävään dataan, on aika siirtyä luomaan visualisointia. Ohjelma perustuu kolmeen erilaiseen sivupohjaan, joiden avulla esityskokonaisuuksia voidaan rakentaa. Nämä ovat: Worksheet, Dashboard ja Storyboard.

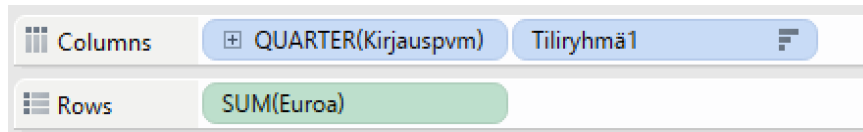
Worksheet on ympäristö, jossa yksittäiset visualisoinnit varsinaisesti rakennetaan ja tyylitellään. Tätä ympäristöä käsitellään ensimmäisenä, koska siitä kokonaisuuden rakentaminen aloitetaan.

Dashboard on koostesivu, jossa yksittäisistä kaavioista voidaan koota isompi kokonaisuus. Dashboard sivuille voidaan sijoittaa useita visualisointeja sekä muuta mediaa kuten tekstiä ja kuvia.

Storyboard on monisivuinen esitystapa visualisointikokonaisuuksille. Tyyliältään se muistuttaa hieman Powerpoint-esitystä. Sen avulla voidaan luoda monisivuisia kokonaisuuksia, jotka ovat valmiita esitettäväksi sellaisenaan.

5.6 Visualisoinnin rakentaminen - Worksheet

Visualisointien luominen aloitetaan sijoittamalla siihen halutut dataelementit. Tässä on kaksi mahdollisuutta, sijoittaa dataa pysty- ja vaakakseleille perinteiseen tapaan tai akselien sijaan sijoittaa dataa vain esim. Koko- ja värimuuttujiin.



Kuva 10. Dataelementtejä sijoitettu molemmille akseleille.

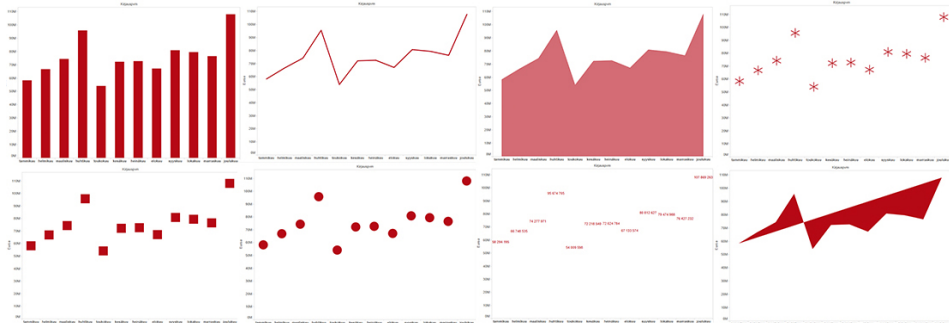
Sitä mukaa kun dataa lisätään, ohjelma muodostaa alustavasti sopivimman mallin, jotta käyttäjä pystyy havainnollistamaan paremmin lisäämänsä elementit. Kun dataelementtejä on lisätty, ”Show me” valikko kertoo mitä erilaisia malleja on mahdollista luoda käyttäjän valitsemasta datasta.

Valikkoon on tallennettu mallipohjia erilaisista visualisoinneista ja sen avulla voidaan nopeasti tarkastella miltä erilaiset mallit näyttäisivät valitun datan kanssa. Valikko myös ilmoittaa minkä tyyppisiä dataelementtejä mikäkin malli tarvitsee, jotta se on mahdollista luoda.



Kuva 11. Show me – valikko. Ohjelma ilmoittaa, että viivakaavioon tarvitaan yksi kvantitatiivinen-elementti ja yksi kvalitatiivinen-elementti, jotta se voidaan koostaa.

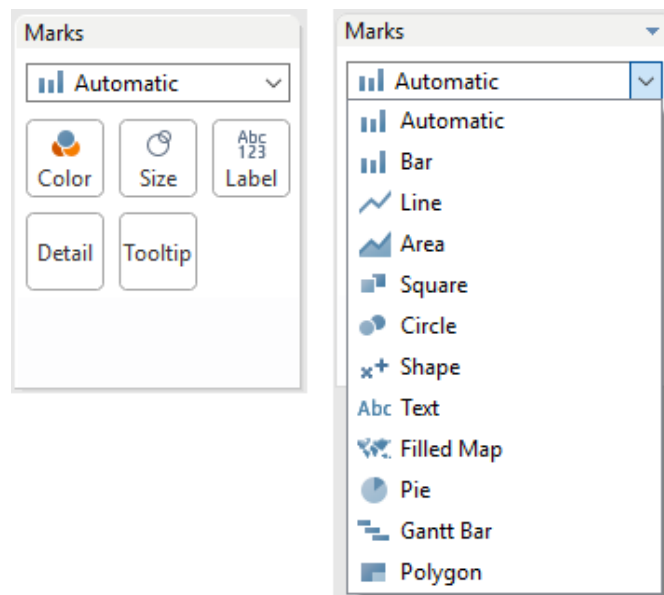
”Marks” paneeli pitää sisällään keskeisimpiä toimintoja visualisoinnin muotoiluun. Siitä voidaan määritellä niin elementtien värejä, kokoa, muotoilla infolaatikon ja selitteiden tekstejä ja muutamia muita toimintoja. Suurin osa visualisoinnin varsinaisesta rungosta määritellään tämän paneelin avulla.



Kuva 12. Erilaisia kuvaajaelementtejä

Ensimmäisenä ylhäällä on vetovalikko, josta valitaan haluttu grafiikkaelementti tyyppi. Valittavat tyypit ovat: palkki, viiva, alue, neliö, ympyrä, muut muodot, teksti, kartta, piirakka, Gantt-viivat, ja polygoni. Näistä muutamia esitelty kuvassa 11.

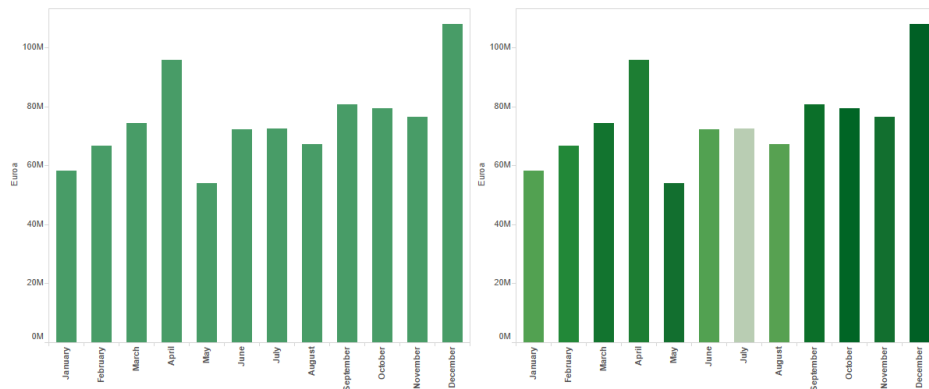
Muodot toistuvat erilaisilla riippuen siitä onko käytössä akselillinen kaavio vai ei, joten käytännössä vaikka ohjelma antaa aina valita minkä tahansa vaihtoehtoista, eivät ne välttämättä joka tilanteessa toistu oikein.



Kuva 13. ”Marks”-paneeli ja vetovalikon sisältö.

Elementtien väriä voidaan muuttaa ihan vain kosmeettisista syistä, mutta väriin voidaan myös sitoa dataa. Kuvassa 13 vasemmalla näkyvä kaavio kuvaa ostojen kokonaismäärää eri kuukausilla euroissa vuonna 2014. oikeanpuoleisessa kaaviossa on otettu mukaan ostojen lukumäärä, joka on sidottu värimuuttujaan ja ilmaisee tummemmalla sävyllä isompaa ostojen

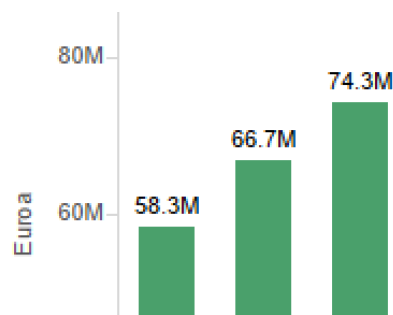
lukumäärä. Jos värimuuttujaan sitoo ei-laskennallisen arvon, sillä voidaan perinteiseen tapaan ryhmitellä ja kategorisoida arvoja.



Kuva 14. Ostojen lukumäärä määritetty värimuuttujaan.

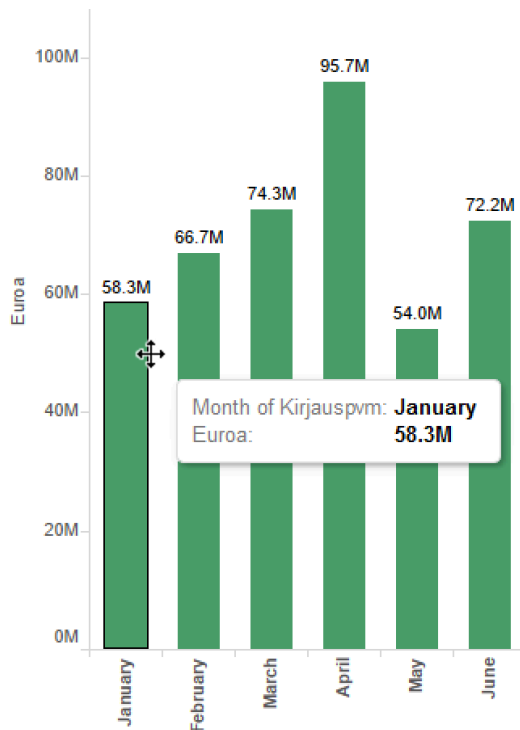
Samalla tavalla kuin värimuuttuja, toimii myös kokomuuttuja. Mikäli kuvan 10 kaavioon lisättäisiin ostojen lukumäärä käyttäen kokomuuttujaa värin sijaan, palkkien paksuus ilmaisisi tällöin lukumäärää. Tämän tyyppisessä pylväskaaviossa väri- tai kokomuuttujan käyttäminen äsken mainitulla tavalla saattaa olla melko epäselvä ja vaikeaselkoinen ratkaisu. Ne toimivat huomattavasti paremmin monissa muunlaisissa malleissa kuten karttaissa.

Muita ”Marks” paneelin toimintoja on Label eli etikettitekstitoiminto. Sen avulla voidaan lisätä dataelementin arvo tekstinä suoraan malliin. Kyseinen toiminto on perusominaisuus ja tuttu esimerkiksi Excelistä. Se on toimiva ratkaisu, mikäli on tarkat luumäärät ovat tärkeitä ja akselin asteikko on suuri, jolloin siitä on vaikea nähdä tarkkaa arvoa. Kuvassa 14. Esimerkki etikettitekstistä käytännössä. Lukujen pyöristyssääntöjä voidaan muokata Format-paneelin kautta, joka esitellään tulevissa kappaleissa.



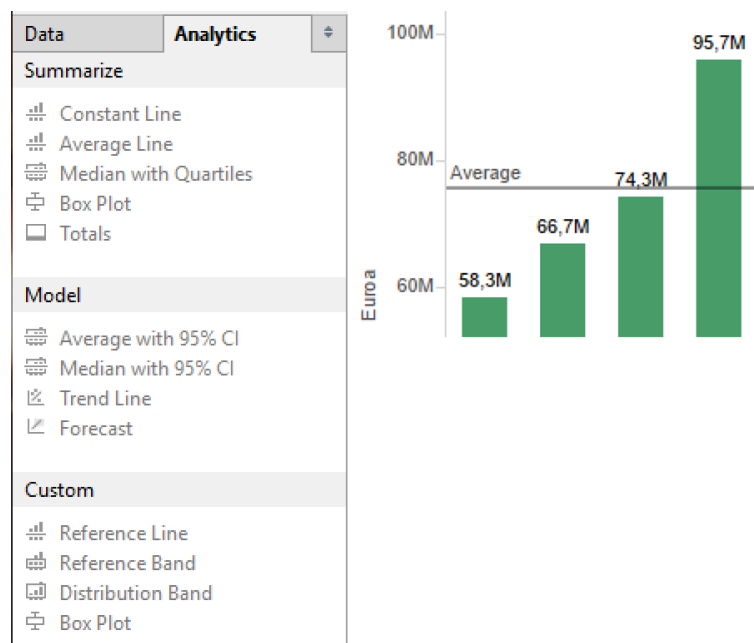
Kuva 15. Label-teksti palkin päällä. Arvo pyöristetty M€ yhdellä desimaalilla.

Seuraavana toimintona on Infoikkuna. Ikkuna kertoo yksityiskohtia tietyttä valitusta dataelementistä. Tableau automaattisesti lisää kaikki akseleilla lisätyt elementit myös infolaatikkoon mutta sen sisältöä voidaan täysin vapaasti muotoilla halutun laiseksi. Kuten etikettitekstiäkin, voidaan infoikkunan tekstiä muotoilla samalla tekstieditorilla joka pitää sisällään perus tekstinkäsittelytoiminnot.

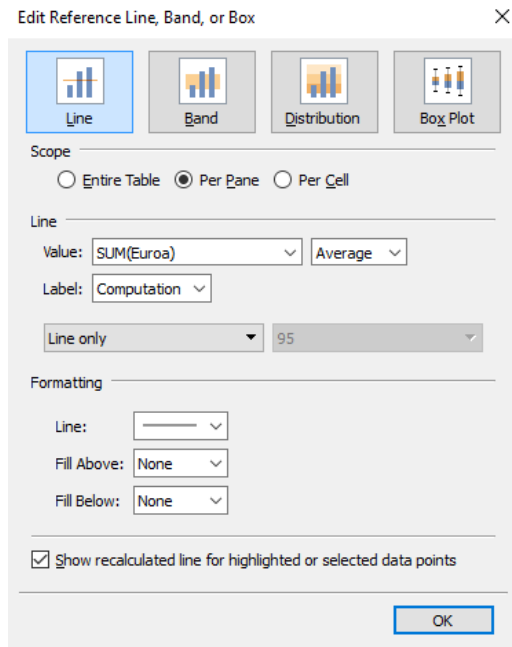


Kuva 16. Infoikkuna

Analytics-paneelin avulla voidaan kuvaajaan piirtää erilaisia laskennallisia funktioita viivoina esitettyinä. Esimerkiksi keskiarvon, mediaaniarvon ja erilaisten ennuste-viivojen lisääminen onnistuu tästä paneelista. Nämä toiminnot ovat käytettävissä vain, mikäli visualisoinnissa on akseleille sijoitettua dataa. Viivojen käyttäytymistä ja esitystapaa voidaan muotoilla niiden oman edit-valikon kautta, joka näkyy kuvassa 17.



Kuva 17. Keskiarvoviiva käytössä.

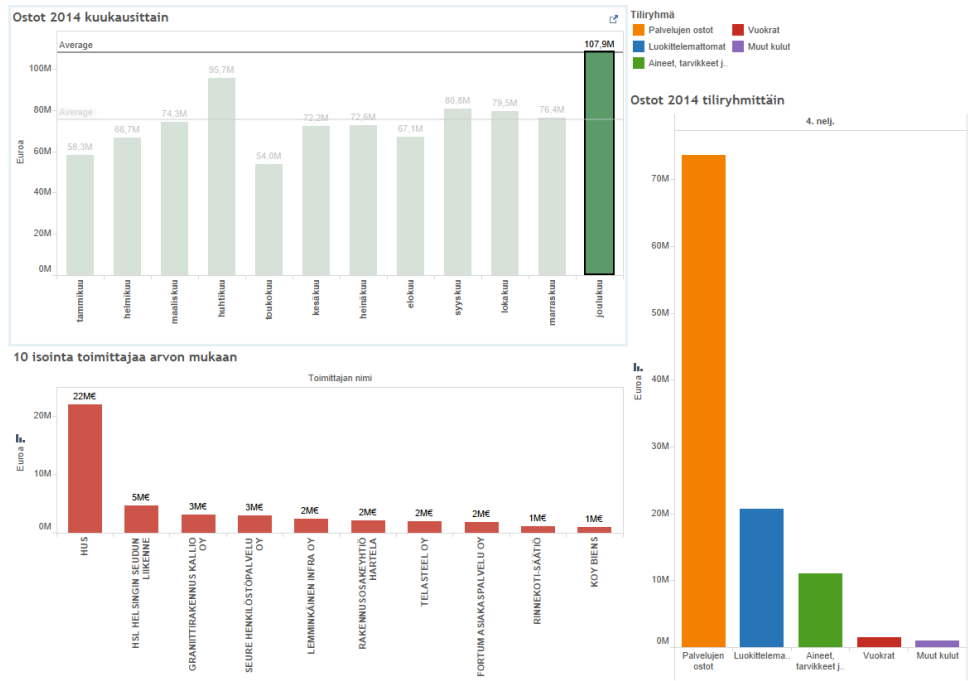


Kuva 18. Viivojen editointi-ikkuna.

Tämän paneelin avulla voidaan määrittellä esityksen graafisia yksityiskoh-
tia kuten värejä ja fontteja. Paneeli on jaettu viiteen eri välilehteen, joissa
määritellään fontit, tekstinasettelu, väritys, reunukset ja apuviivat. Näistä
jokaisessa voidaan muutokset kohdistaa joko koko arkkiin tai vain riveihin
tai sarakkeisiin.

Akseleiden asetuksia määritetään omasta valikostaan. Tässä paneelissa
voidaan muokata akselin skaalaa sekä määrittellä apuviivojen tiheys ja
käyttäytyminen sekä asettaa akselin otsikkoteksti.

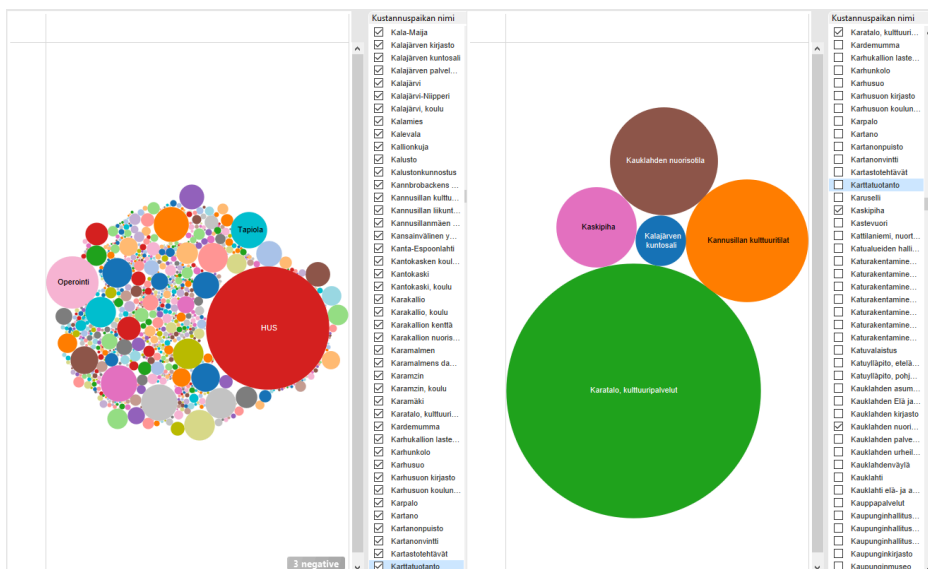
Dashboard sivulla jossa on useita kuvaajia, voidaan kuvaaja määrittää
toimimaan myös suodattamina. Käyttäjä voi valita dataelementin jolloin se
aktivoituu suodattimeksi ja suodattaa myös muiden samalla Dashboardissa
olevien kuvaajien datan valitulla tavalla.



Kuva 19. suodattimeksi valittuna ostot/kk kuvaajasta joulukuuta, jolloin se rajaa kahden muunkin kuvaajan esittämään vain joulukuulla olleita tapahtumia. Kuvaaja pitää ensin määrittää ”use a a filter” tilaan.

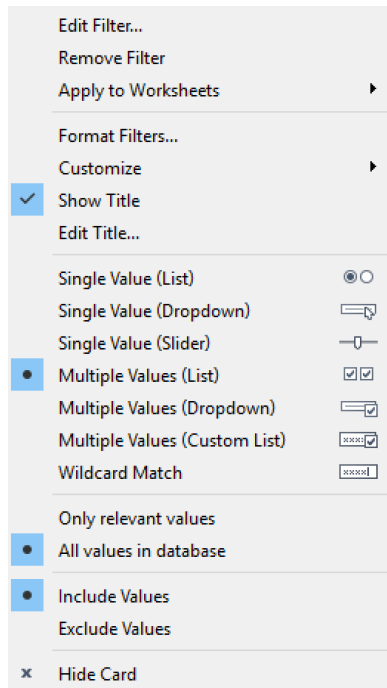
Kuvassa 20 on valittu suodattimeksi kustannuspaikan nimi. Vasemman puoleisessa esimerkissä on valittu kaikki ja oikean puoleisessa vain muutama. Kun dataelementtiä halutaan suodattaa, se vedetään ”Filter” laatikkoon, jonka jälkeen suodatuksen asetuksia voidaan muokata. Suodatettavia elementtejä voi olla useita ja niille voidaan määrittää erilaisia ehtoja sekä valmiita settejä.

Setit ovat tallennettuja valintojen sarjoja jolloin voidaan tallentaa valmiiksi tietyt tilanteet, jotta katsojan on helpompi siirtyä nopeasti tilanteesta toiseen, ilman että hänen täytyy aina valita käsin halutut arvot.



Kuva 20. Kun rajataan vertailtavia kohteita, on vertaaminen huomattavasti helpompaa.

Myös suodatuksen valintatauluun on muutamia eri vaihtoehtoja. Riippuen suodatetun datan tyypistä voi olla tarpeellista käyttää joko monivalintaa tai yksivalintaa ja näille molemmilla voidaan valita muutamista erilaisesta valikkomallista. Valikko näkyy kuvassa 21.



Kuva 21. Suodatusvalikon muotoiluasetuksia

Suodatusvalikossa tehdyt asetuksen vaikuttavat loppukäyttäjän näkymään ja näin voidaankin rajoittaa suodattimen toiminnallisuutta, kuten tarpeelliseksi nähdään. Esimerkiksi luettelotyypin valinta määrittää voidaanko suodattaa useita arvoja isommasta joukosta, vaiko tarkastella vain yhtä arvoa kerrallaan. Suodatuksen voi myös asettaa käänteiseksi, jolloin kaikki muut paitsi valittu tai valitut arvot ovat näkyvissä. Myös suodatinikkunan ulkonäköä ja otsikkoa voidaan jonkin verran muotoilla.

Pages toiminnon avulla voidaan näyttää vain tietty osa datasta kerrallaan. Sen toiminnallisuus on tietyllä tapaa melko lähellä suodattamista, mutta se tarjoaa erilaisia toimintoa suodattamiseen verrattuna. Yhtenä isoimpana erona suodattamiseen on että pages-toiminnolla rajattu aineisto säilyttää akselien skaalan kun taas suodattamista käytettäessä, muuttuvat akselit esittämään suodatetun aineiston mukana.

5.6.1 Akselittomat mallit

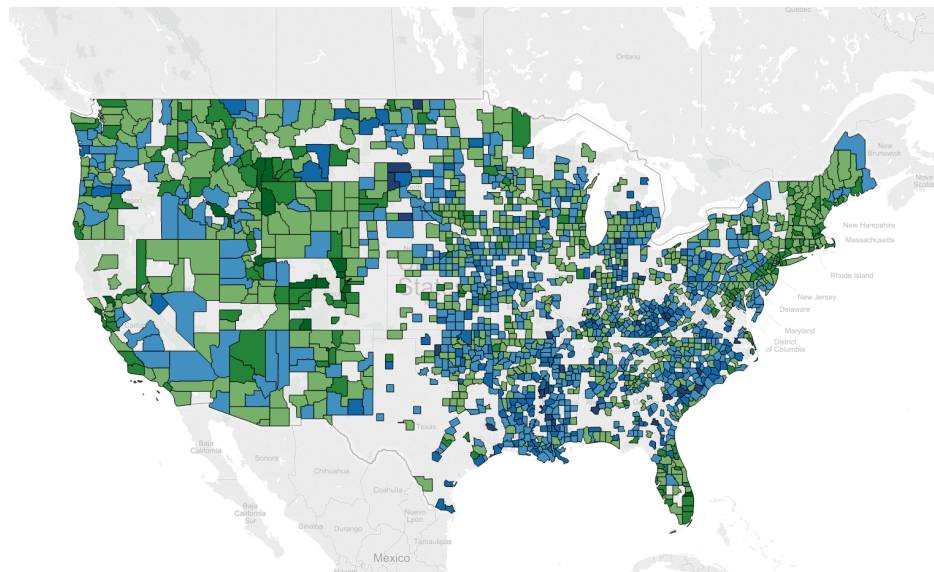
Kuten aiemmin mainittiin, voi visualisointia lähteä rakentamaan myös sijoittamatta dataa suoraan akseleille. Tämän tyyppisessä mallissa yleensä arvot ovat yksittäisiä grafiikkaelementtejä, joissa määrää ja ryhmittelyä kuvataan väri- ja kokomuuttujilla. Tämän tyyppisillä malleilla saadaan hyvin vaihtelua perinteisiin malleihin ja niillä voidaan monissa tapauksissa havainnollistaa dataa huomattavasti paremmin kuin perinteisillä malleilla.

Edellisessä kappaleen kuvassa 19 näkynyt pallovisualisointi on hyvä esimerkki akselittomasta visualisoinnista. Siinä pallot kuvaavat eri toimijoita, joilta Espoon kaupunki on tehnyt ostoja. Pallojen koko kuvaa tässä tapauksessa jokaisen toimijan tilausten yhteenlaskettua euromääräistä arvoa, kun taas pallojen väri kuvaa minkä alan toimia on kyseessä.

5.6.2 Karttapohjaiset mallit

Karttapohjaisia malleja voidaan käyttää, kun käytetty data sisältää jonkinlaisen paikkatiedon. Paikkatieto voi olla koordinaattitietoa tai pelkästään kaupungin tai paikan nimiä. Nämä mallit perustuvat karttakuvaan, jonka päälle sijoitetaan värjättyjä alueita tai symboleja. Näin voidaan tehokkaasti kuvata jonkin asian määrää tai tiheyttä tietyllä maantieteellisellä alueella.

Tableausta löytyy sisäänrakennettu karttapohja, jota voidaan käyttää paikkatietovisualisointien kanssa.



Kuva 22. Tableaun esimerkissä kuvattuna USA:n kansalaisten liikuntatottumuksia.

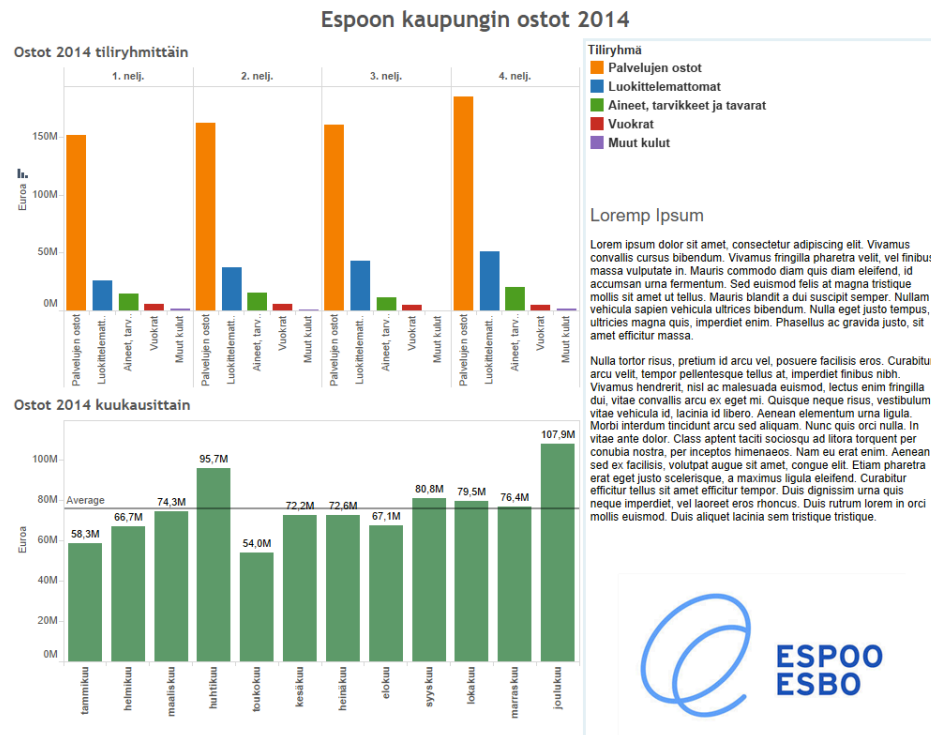
5.7 Koostaminen, Dashboard

Kun esityksen kannalta tärkeät mallit on luotu, niistä voidaan rakentaa koostesivuja. Koosteessa luodaan yksittäinen näkymä joka voi sisältää luotujen kaavioiden lisäksi muita elementtejä kuten tekstiä, kuvia ja web-upotuksia.

Sisältöä sivulle lisättäessä voidaan määritellä, onko lisättävä kappale ”tiled” tilassa, jolloin se ryhmittyy muiden elementtien kanssa omaksi laatikokseen, vaiko ”floating” eli kelluvassa tilassa jolloin sitä voidaan vapaasti liikutella muiden objektien yläpuolella ja tarvittaessa asemoida se jonkin toisen objektin päälle.

Dashboard sivua tehdessä pitää sivulle myös asettaa sen koko. Tableaussa on tallennettu valmiita esi-asetuksia erilaisille alustoille kuten puhelimen

tai tabletin ruudulle sekä myös valmiita arkkikokoja paperille jolloin voidaan tarvittaessa luoda valmiiksi sovitettut versiot eri käyttökohteita varten.



Kuva 23. Koostesivu johon sijoitettuna kolme eri elementtiä.

5.8 Esitykset, Storyboard

Kuten nimi kertoo, on storyboard tarkoitettu isompaan kokonaisuuteen tai esitykseen, jossa datalla havainnollistetaan isompaa kokonaisuutta tai kerrotaan tarinaa. Se on monisivuinen kokonaisuus, joka perus ajatukseltaan muistuttaa hieman Powerpoint-esitystä. Se koostetaan Dashboard tai worksheet sivuista.

Storyboard on esityskokonaisuus, jonka avulla voidaan esittää jokin asia kokonaisuudessaan tai ”kertoa tarinaa” datan avulla. Se perustuu erillisiin sivuihin tai välilehtiin joihin valmiita Worksheet tai Dashboard sivuja voidaan sijoittaa. Storyboard ei tarjoa enää erillisiä muotoilutyökaluja, ja siinä voikin ainoastaan vain lisätä tekstilaatikoita, nimetä sivut valintapalkkiin ja lisätä otsikon.

6 VISUALISOINNIN TALLENTAMINEN JA JULKAISU

Kun luotu visualisointikokonaisuus on valmis ja se halutaan julkaista, sen jakamiseen on muutama eri vaihtoehto. Tiedosto voidaan tallentaa Tableau omassa projektitiedostoformaattissa, jolloin sitä voidaan katsella vain Tableau viewer ohjelmaa käyttäen. Tämä vaihtoehto ei ole erityisen kätevä ja helppo tapa esityksen jakamiseen. Tiedosto on ensin lähetettävä vastaanottajalle, minkä jälkeen hänen on ladattava katseluohjelma. Tämä tapa onkin lähinnä hyvä vain sisäiseen käyttöön, jossa malli lähetetään edelleen jatkotyöstettäväksi toiselle henkilölle.

Visualisointi voidaan myös tallentaa tavalliseksi PNG tai JPEG kuvatiedostoksi jolloin se voidaan helposti julkaista verkossa, mutta tällöin menetetään luonnollisesti kaikki Tableau interaktiiviset toiminnot, joita esitykseen on rakennettu. Mikäli visualisointi on isompi kokonaisuus joka pitää sisällään monia malleja, sen julkaiseminen erillisinä kuvina voi olla myös melko työlästä.

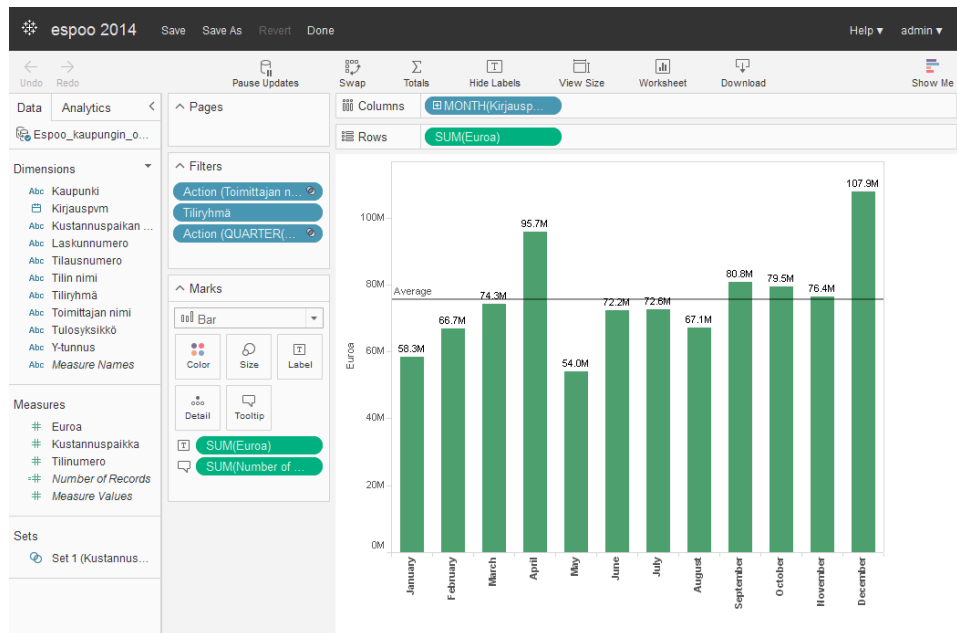
Kuitenkin kaikista käyttökelpoisin vaihtoehto johon ohjelma on pääsääntöisesti tarkoitettukin, on jakaa visualisoinnit verkossa käyttäen joko omaa Tableau server-palvelinta tai Tableau online SaaS-palvelua. Tällöin visualisoinnit säilyttävät interaktiivisuutensa täydellisesti ja niitä voidaan tarkastella verkkoselaimella käytännössä täysin sellaisina kuin ne näkyvät Tableau desktop ohjelmistossakin. Olipa käytössä sitten oma palvelin tai Online-palvelu, voidaan visualisointien näkyvyyttä ja jakamisasetuksia muuttaa halutun laisiksi. Tekniikaltaan ne pohjautuvat Ajax pohjaiseen Javascriptiin, joten niiden käyttäminen ei vaadi mitään kolmansien osapuolten selainlisäosia joista esimerkkinä voidaan mainita Adobe Flash tai Microsoft Silverlight. Tämän ansiosta niiden käyttäminen mobiililaitteella pelkästään selaimen avulla on mahdollista.

6.1 Tableau server

Online –ja server-ohjelmistoihin visualisointiesitykset siirtyvät suoraan työpöytä sovelluksesta, joten käsin tiedostoja ei tarvitse missään tilanteessa käsitellä. Kun esityskokonaisuus tai halutut osat siitä on siirretty, Server-ohjelmisto tarjoaa paljon erilaisia toimintoja liittyen visualisoinnin käsittelyyn ja jakamiseen.

Palvelimella voidaan määrittää eri käyttäjiä ja käyttäjien luku ja muutosoikeuksia palvelimelle ladattuihin visualisointeihin. Tähän voidaan käyttää palvelimen oman paikallisen käyttäjähallinnan sijaan myös Windows Active Directory –hakemistoa mikäli käytössä on sitä kautta hoidettu yleinen käyttäjäoikeuksien hallinta

Visualisointimalleja on mahdollista myös muokata edelleen palvelimella. Selaimessa toimiva editori ei luonnollisesti tarjoa kaikkia työpöytä ohjelman toimintoja, mutta kuitenkin on toiminnaltaan melko laaja. Kuvassa 24 on esitettynä visualisoinnin editointinäkökulma palvelimelta käsin, joka muistuttaa melko paljon työpöytäohjelman näkymää.



Kuva 25. Editointinäkömä palvelimella.

Palvelimelta on myös mahdollista saada upotuskoodi visualisointeihin, jolloin niitä voidaan sijoittaa esim. Verkkosivuille. Esimerkiksi jokin moniosainen Dashboard voitaisiin laittaa suoraan yrityksen intranettiin tarkasteltavaksi.

Palvelin voi toimia visualisoinneissa käytettävän datan kanssa kahdella tavalla. Ottamalla niistä paikallisen kopion itselleen tai olemalla suoraan yhteydessä tietokantaan. Tämä määrittää työpöytä ohjelmistossa, kun käytettyä datalähdettä ollaan tuomassa, mutta asetusta voidaan kuitenkin muuttaa myös palvelimelta käsin.

Paikallisen datakopion hyviä puolia on se, että mikäli useampi visualisointi hyödyntää samaa dataa, voidaan ne määrittää käyttämään samaa paikallista datakopiota sen sijaan että jokainen malli olisi erikseen yhteydessä ulkoiseen tietokantaan hakeakseen saman datan uudestaan ja uudestaan. Tämän tyyppisessä tapauksessa paikallinen datakopiointi vähentää huomattavasti turhaa verkkoliikennettä. Palvelimelle voidaan määrittää aika-tiloja datakopioiden hakemiseen, ja eri datojen päivitys-syklejä voidaan priorisoida eri tasoille, jolloin voidaan välttää palvelimen ylikuormittuminen.

Toisena vaihtoehtona on live-yhteys, jolloin paikallista kopio datasta ei oteta, vaan data haetaan aina suoraan palvelimelta, kun visualisointi avataan tarkastelua varten. Live yhteyksien liittyen voidaan palvelimelle määrittää cache-asetus eli välimuisti pois päältä, jolloin tarkasteltava data on aina varmasti tuoreinta mahdollista. Tosin jos cache-toimintoa ei käytetä, se vaikuttaa negatiivisesti palvelimen nopeuteen koska palvelin noutaa aina uusimman datan jokaisella katselukerralla.

6.2 Server API

Palvelimella on myös käytettävissä oma Javascript rajapinta, jonka avulla visualisointeja voidaan integroida erilaisiin web-sovelluksiin. Sen avulla on mahdollista liittää sivulla olevia HTML elementtejä kuten liukusäätimiä ja valintalaatikoita toimimaan yhdessä visualisoinnin kanssa. Tällöin visualisoinnin kontrollit voidaan esimerkiksi liittää samoihin ohjain-elementteihin, joiden avulla ohjataan sivun muita mahdollisia toiminnallisuuksia.

6.3 Tableau Online

Kolmantena vaihtoehtona on käyttää Tableau online palvelua ja siirtää visualisointi sinne. Tableau online on käytännössä Tableau server tarjottuna SaaS (software as a service) mallisena palveluna. Se on suurimmaksi osaksi toiminnaltaan samanlainen kuin Server-ohjelmistokin, mutta muutamia poikkeuksia löytyy:

Yhtenä isona eroavaisuutena on hinta. Pidemmällä aikavälillä online palvelu tulee huomattavasti kalliimmaksi käyttää kuin oma palvelin Server-ohjelmistolla. Toki palvelimen ylläpito ja toiminta on täysin Tableaun vastuulla, joten käyttäjän ei tarvitse huolehtia ylläpidollisista asioista lainkaan.

Tableau online tukee myös live-päivitystä kuten Server-ohjelmistokin, mutta jonkin verran suppeampana. Oline pystyy yhdistämään vain pilvipalveluissa sijaitseviin tietokantoihin, joten esimerkiksi omalla palvelimellaan ylläpidettyyn MySQL-tietokantaan yhdistäminen ei onnistu. Kirjoitus hetkellä tuettuja palveluita ovat mm. Google Cloud SQL, Amazon Redshift ja Azure SQL Data Warehouse.

Suppeamman tietokantaan yhdistämismahdollisuuksien vuoksi voi olla pakko käyttää dataa extract-tilassa. Tällöin käytettävä data on ajettava Tableaun omilla palvelimilla, jolloin pitää miettiä onko visualisoinnissa oleva data arkaluontoista ja voidaanko sitä siirtää ulkopuoliselle palvelimelle. Online

Haastattelussa Nylund kuitenkin mainitsi, että todennäköisesti data mitä he Tableaulla käyttäisivät, on sellaista joka ei sisällä mitään kovinkaan arkaluontoista. Tällöin datan lähettäminen USA:ssa sijaitseville palvelimille ei muodostu ongelmaksi. (Kukkamäki & Nylund, haastattelu 18.1.2016)

7 TUTKIMUKSEN TULOKSET

Tutkimuksen tarkoituksena oli tutkia, mitä datavisualisointi itsessään tarkoittaa, millä ja miten data muutetaan visuaaliseen muotoonsa oikeaoppisesti, sekä myös minkälainen työkalu Tableau on verkkopohjaisten datavisualisointien luomiseen. Työn pohjalta toimeksiantaja eli Älykkäät palvelut –tutkimusyksikkö voi hahmottaa ohjelmiston soveltuvuutta heidän käyttöönsä ja tämän jälkeen harkita ohjelman tuotantokäyttöön ottamista.

Teoriaosuus pyrki vastaamaan kysymykseen: mitä tarkoittaa datavisualisointi ja mitä tulee huomioda sellaista luodessa. Datan visuaalisen muotoon muuttaminen antaa ihmiselle mahdollisuuden hahmottaa suuria määriä monimutkaista dataa. Teoriaosuudessa käytiin läpi asioita, joita tulee tiedostaa, kun datalle ollaan antamassa visuaalinen ulkoasu, mikä on oikein ja mikä väärin.

Lopputuloksena voidaan todeta, että Tableau on hyvin kattava ohjelmisto ja siinä on laaja tuki erilaisille tietolähteille ja sen on toimiva niin pienten kuin suurtenkin datamäärien kanssa. Se on joustava ja mukautuva työkalu, joka tarjoaa mahdollisuuden luoda tehokkaasti näyttäviä kaavioita ja visualisointeja perinteiseen tapaan staattisena kuvana kuin myös laajoja esityskokonaisuuksia verkkoon. Se pitää sisällään käytännössä kaikki toimeksiantajan vaatimat ominaisuudet ja se myös integroituu kaikkiin tällä hetkellä toimeksiantajalla käytössä oleviin tietokantoihin ja lähteisiin.

Javascript-rajapintansa avulla Tableaulla luotuja datavisualisointeja voidaan integroida erilaisilla tavoilla verkkosivuille ja erilaisiin verkkosovelluksiin. Koska visualisointien tekniikka perustuu Javascriptiin, ne toimivat hyvin myös mobiililaitteilta käytettäessä.

Ohjelmistoperhe kokonaisuudessaan on melko laaja paketti, jonka palvelin-puoli tarjoaa tehokkaan väylän esitysten jakamiseen verkossa niin rajatusti organisaation sisällä kuin myös vapaasti ulkopuolelle. Palvelimen käyttäjähallintaominaisuuksien vuoksi se mahdollistaa myös monen ihmisen työskentelyn dokumentteihin verkon yli organisaation sisällä. Mahdollisuus käyttää ulkoista Active Directory-pohjaista käyttäjähallintaa helpottaa ohjelman käyttöönottoa suuremmissa organisaatioissa.

Tableauhun tutustuminen oli mielenkiintoinen kokemus ja oikeastaan yllätyin ohjelman helppokäyttöisyydestä ja loogisuudesta. Ohjelmisto on mielestäni helposti lähestyttävästi käyttöliittymästä huolimatta hyvin laaja ja pitää sisällään suuren määrän ominaisuuksia. Tämän opinnäytetyön puitteissa ei perehdytty erityisen syvälle ohjelman kaikkiin ominaisuuksiin, mutta mielestäni ohjelmistolla on potentiaalia ja joustavuutta sopia monen erilaiseen rooliin.

Opinnäytetyön tekemisessä suurimpana ongelmana oli aiheen liian laaja rajaus, jota jouduin toistuvasti kaventamaan. Aihe on hyvin mielenkiintoinen ja laaja ja siitä tietyn osa-alueen valitseminen tuotti itselleni jonkin verran vaikeuksia.

LÄHTEET

Creativyst. n.d-a. The comma separated value (CSV) file format. Viitattu 12.7.2015.

<http://creativyst.com/Doc/Articles/CSV/CSV01.html>

Creativebloq. 2013. The science behind data visualization. Viitattu 8.9.2015.

<http://www.creativebloq.com/design/science-behind-data-visualisation-8135496>

Datalabs. n-d.a. 15 most common types of datavisualizations. Viitattu 8.9.2015.

<http://www.datalabs.com.au/articles/15-most-common-types-of-data-visualisation/>

Haasio, A & Vakkari, P. n-d.a. Viestintätieteiden yliopistoverkoston opimateriaalit, informaatiotutkimuksen perusteet. Viitattu 22.9.2015.

<https://viestintatieteet-wiki.wikispaces.com/Informaatiotutkimuksen+perusteet>

Hämeen ammattikorkeakoulu. n.d-a. Älykkäät palvelut – tutkimusyksikkö. Viitattu 20.6.2015

<http://www.hamk.fi/tyoelamalle/tutkimusyksikot/alykkaat-palvelut/Sivut/default.aspx>

Hämeen ammattikorkeakoulu n.d-b. Yksiköt. Viitattu 30.6.2015

<http://www.hamk.fi/tietoa-hamkista/yksikot/Sivut/default.aspx>

Hämeen ammattikorkeakoulu. n.d-c. Hakijalle. Viitattu 8.7.2015

<http://www.hamk.fi/hakijalle/Sivut/default.aspx>

informaatiomuotoilu.fi. 2012. Pitääkö datavisualisoinnin olla kaunis?. Viitattu 12.9.2015

<http://informaatiomuotoilu.fi/2012/06/pitaako-visualisoinnin-olla-kaunis/>

Interaction desing foundation. n.d-a. Data visualization for human perception. Viitattu 8.9.2015.

https://www.interaction-design.org/encyclopedia/data_visualization_for_human_perception.html

Jokela, M. 2011. Mitä tieto on? Viitattu 8.4.2016

http://inside-the-plm.blogspot.fi/p/mita-tieto-on_12.html

Jyväskylän yliopisto. 2004. Hahmolait käytettävyyden parantajina. Viitattu 18.9.2015.

<http://www.mit.jyu.fi/opetus/opinnayte/LuK/Hahmolait/>

Kettunen, E. 2015. Kunnat.net. Datavisualisointi muuntaa tiedon helposti omaksuttavaksi. Viitattu 30.6.105.

<http://www.kunnat.net/fi/tietopankit/kolumnit/2015/Sivut/datavisualisointi-muuntaa-tiedon-helposti-omaksuttavaksi.aspx>

Kuusela, V. 2000. Tilastografiikan perusteet. Helsinki. Oy Edita Ab.

Lengstorg, J. n.d-a. Copterlabs.com. JSON: what it is, how it works, & how to use it. Viitattu 9.7.2015.

<http://www.copterlabs.com/blog/json-what-it-is-how-it-works-how-to-use-it/>

Liikenne –ja viestintäministeriö. 2013. Big data suomessa. Viitattu 18.9.2015.

http://www.lvm.fi/docs/fi/2497123_DLFE-21601.pdf

Murray, S.2013. Interactive data visualization for the web. Sebastopol, CA, USA: O'Reilly Media.

Padmanabhan, U. 2013. How to make data visualizations better with gestalt laws. Viitattu 3.1.2016

<http://sixrevisions.com/usability/data-visualization-gestalt-laws/>

Sarja, J. 2006. Verkkopedagogi.net, Relaatiokanta. Viitattu 22.9.2015.

<http://verkkopedagogi.net/vanhat/fi/sisalto/materiaalit/access2003/luku0375c6.html?C>

Sharma, N. 2015. Thenextweb.com. The 14 best data visualization tools. Viitattu 12.7.2015.

<http://thenextweb.com/dd/2015/04/21/the-14-best-data-visualization-tools/>

Skau, D. 2012. How to use maps in data visualization. Viitattu 3.1.2016

<http://blog.visual.ly/you-are-here-using-maps-in-data-visualization/>

Tampereen yliopisto. n.d-a. Visualisointi. Viitattu 6.7.2015

<https://blogs.uta.fi/datajournalismi/visualisointi/>

Tableau. n.d-a. Who's using Tableau? Viitattu 31.8.2015

<https://www.tableau.com/about/customers>

Tableau. n.d-b. History in the making. Viitattu 4.1.2016

<https://careers.tableau.com/ourstory>

Tableau. n.d-c. Tableau desktop. Viitattu 16.4.2016

<http://www.tableau.com/products/desktop>

Tableau. n.d-d.. Tableau online. Viitattu 16.4.2016

<http://www.tableau.com/products/cloud-bi>

Tableau. n.d-e. Tableau Server. Viitattu 16.4.2016

<http://www.tableau.com/products/server>

Tableau. n.d-f. Tableau reader. Viitattu 16.4.2016
<http://www.tableau.com/products/reader>

Tableau. n.d-g. Buy Tableau. Viitattu 9.5.2016
<https://buy.tableau.com>

University of Saskatchewan. 2010. Useful junk? Effects of visual embellishment on comprehension and memorability of charts. Viitattu 12.9.2015
<http://hci.usask.ca/publications/view.php?id=173>

YLE. 2012. Näin presidenttiehdokkaat ja heidän seuraajansa sijoittuvat twitterissä. Viitattu 6.7.2015.
<http://blogit.yle.fi/kehitys-kehittyy/nain-presidenttiehdokkaat-ja-heidan-seuraajansa-sijoittuvat-twitterissa>

YLE. 2015. Vaalipallo on vain yksi tapa seurata tuloslaskentaa. Viitattu 6.7.2015
http://yle.fi/uutiset/vaalipallo_on_vain_yksi_tapa_seurata_tuloslaskentaa/7936686

Älykkäät palvelut –tutkimusyksikkö. 2016. Datankeräsymalli.

HAAASTATTELUT

Kukkamäki, J. 2016. Tutkimuskoordinaattori. Älykkäät palvelut. HAMK. Haastattelu 18.1.2016

Nylund, T. 2016. Tuntiopettaja. HAMK. Haastattelu 18.1.2016

