

# Työhön osallistumisen segmentaatiomalli

TULOKSIA VÄESTÖTASON DATALÄHTÖISESTÄ RYHMITTELYANALYYSISTÄ

Jarno Turunen  
Suvi Ahtinen  
Matti Joensuu

# **Työhön osallistumisen segmentaatiomalli**

Tuloksia väestötason datalähtöisestä  
ryhmittelyanalyysistä

Jarno Turunen, Suvi Ahtinen, Matti Joensuu

Työterveyslaitos

Työterveyslaitos

PL 40

00251 Helsinki

[www.ttl.fi](http://www.ttl.fi)

Toimitus: Jarno Turunen, Suvi Ahtinen, Matti Joensuu

Piirrokset: Suvi Ahtinen

© 2022 Työterveyslaitos ja kirjoittajat

Tämän teoksen osittainenkin kopiointi on tekijänoikeuslain (404/61, siihen myöhemmin tehtyine muutoksineen) mukaisesti kielletty ilman asianmukaista lupaa.

ISBN 978-952-391-061-4 (pdf)

## Tiivistelmä

Tämän tutkimuksen tavoitteena oli tunnistaa työkyvyltään toisistaan eroavia ryhmiä työikäisestä väestöstä. Ryhmien tunnistamisessa hallinnollista rekisteritietoa työssäolosta ja erityisesti työkyvyttömyyteen liittyvistä etuuksista vuosilta 2010-2014 noin 90 % väestötöksestä. Ryhmittely toteutettiin hyödyntäen aineistolähtöisiä ryhmittelymenetelmiä.

Työkyvyttömyyteen liittyviä etuuksia saaneiden henkilöiden ryhmiä oli yhdeksän erilaista. Ryhmiin kuului kaiken kaikkiaan yli 600 000 työikäistä suomalaista. Ryhmät erosivat toisistaan työssäolo- ja työttömyyskuukausien, useiden taustatekijöiden ja saatujen etuuksien osalta. Tässä raportissa on käsitelty ryhmien piirteitä ja ryhmien välisiä eroja.

Aineistolähtöisillä ryhmittelymenetelmillä voidaan muodostaa jäsennettyä tietoa väestöstä ja väestön jakautumisesta erilaisiin ryhmiin. Tätä tietoa voidaan hyödyntää näitä ryhmiä koskevan päätöksenteon suunnittelun ja arvioinnin tukena. Raportissa on tarjottu esimerkkejä tiedon hyödyntämisen mahdollisuuksista ja muodostettu myös skenaariolaskelma työkyvyttömyyteen liittyviä etuuksia saaneiden työvoimanlisäyspotentiaalista.

Tämän tutkimuksen tulokset osoittavat, että ryhmittelyä voitaisiin toteuttaa myös muiden työhön osallistumiseen liittyvien tietojen perusteella. Tällöin muodostuisi tarkentuva kuva tarkastelun kohteena olevasta väestön alaryhmästä.

## Abstract

This study aimed to investigate work disability related subgroups from working-age population of Finland. Clustering was used with administrative register information on working days and work disability benefits of random sample of 90 % of total population for 2010-2014.

A total number of nine groups with received work disability benefits were found. These groups included more than 600 000 individuals that had received work disability benefit of some kind during 2010-2014. They differed with regards to received work disability benefits, other benefits, days worked, and sociodemographic background. This study describes these groups and their differences.

Databased clustering methods can be used to investigate different population samples in more details. The formed information can be used in both ex-ante and ex-post evaluation of different policies and intervention. This study offers an example of using the clustering results in work disability related scenario calculations.

## Sisällys

Tiivistelmä.....	3
Abstract .....	4
Sisällys.....	5
1 Tausta ja tavoite.....	6
2 Aineisto.....	7
2.1 Tutkimuksen tausta-aineisto.....	7
2.2 Tutkimuksessa hyödynnetty aineisto .....	7
3 Menetelmät .....	10
4 Tulokset.....	11
4.1 Taustamuuttajat.....	11
4.2 Ryhmittely koko väestölle.....	13
4.3 Koko väestön ryhmittely ja työssäolokuukaudet.....	24
4.4 Ryhmittely työkyvyttömyyteen liittyviä etuuksia saaneiden väestölle .....	25
5 Pohdinta ja rajoitukset .....	37
5.1 Pohdinta.....	37
5.2 Tulosten hyödynnettävyys.....	38
5.3 Työkyvyttömyyteen liittyviä etuuksia saaneen väestön ryhmittely .....	39
5.4 Skenaariolaskelman hahmottelua.....	42
5.5 Aineistoon liittyviä rajoituksia .....	42
5.6 Menetelmään liittyviä rajoituksia .....	43
6 Seuraavista askeleista .....	44
7 Lopuksi .....	46
8 Lähteet.....	47
9 Liitteet .....	48
9.1 Liitetaulukot.....	48

# 1 Tausta ja tavoite

Väestön ikääntyessä huoltosuhte heikkenee. Huoltosuhteeseen voidaan vaikuttaa suotuisasti lisäämällä työikäisten ja työllisten määrää yhteiskunnassa. Työllisyysasteen nostaminen on ollut tavoitteena Suomessa jo useamman hallituskauden ajan. Se tulee olemaan sitä myös tulevaisuudessa huoltosuhteen heiketessä. Työllisyysasteen trendiluku oli helmikuussa 2022 noin 74,0 prosenttia (Työvoimatutkimus, 2022). Syyskuun 2021 budjettiriihessä julkista taloutta tasapainottavana työllisyysasteen tasona on pidetty 78 prosenttia (Valtioneuvosto, 2021). Tämänhetkinen työllisyysasteen nousu on lisännyt puhetta haasteista työvoiman tarjontapuolella. Toisin sanoen työvoimaa ole enää tarjolla työmarkkinoilla tarjolla oleviin työtehtäviin. Työllisyysasteen nosto ja sen mahdollistava kasvava työvoiman tarjonta edellyttävät suurempaa töihin osallistumisen astetta työikäisessä väestössä.

Työikäisillä on hyvin erilaisia työhön osallistumisen esteitä. Kyselyin toteutetussa Faces of joblessness in Finland -tutkimuksessa (OECD, 2020) löydettiin kahdeksan erilaista heikosti työllistyvien alaryhmää. Työllistymisen edesauttaminen näissä eri ryhmissä edellyttää erilaisia toimia sekä politiikkatasolla että arjessa. Esimerkiksi työkykyiset ja osatyökykyiset työttömät kohtaavat muihin työttömiin verrattuna erilaisia ongelmia pyrkiessään työllistymään.

Yksi mahdollinen ratkaisu työvoiman tarjonnan lisäämiseksi sekä työllisyysasteen nostamiseen on esimerkiksi työkyvyn haasteita omaavien tai ikääntyvien henkilöiden työhön osallistumisen lisääminen. Molemmat ryhmät asettavat uudenlaisia haasteita työikäisten työkyvyn ylläpidolle ja työkyvyttömyyden ehkäisylle. Tässä lyhytraportissa keskitytään kattavasti työkyvyn haasteita omaavaan suomalaiseen työikäiseen väestöön. Tavoitteena on tunnistaa työkyvyn etuuksien käytöltään toisistaan eroavia ryhmiä. Tämä tehdään aineistolähtöisesti käyttäen ryhmittelymenetelmiä. Aineistona toimii kokoelma työikäisen väestön käyttämiin etuuksiin ja työhön osallistumiseen liittyviä, lähes koko työikäisen väestön kattavia hallinnollisia rekistereitä vuosilta 2010-2014.

Tämän tutkimuksen tulosten odotetaan lisäävän tietoa yksilöiden työelämään kiinnittymisen monimuotoisuudesta ja työkyvyn merkityksestä siinä. Tämä voisi edistää yhteiskunnallisten politiikkatoimenpiteiden arviointia erilaisille ryhmille sekä ennen että jälkeen toimenpiteiden.

## 2 Aineisto

### 2.1 Tutkimuksen tausta-aineisto

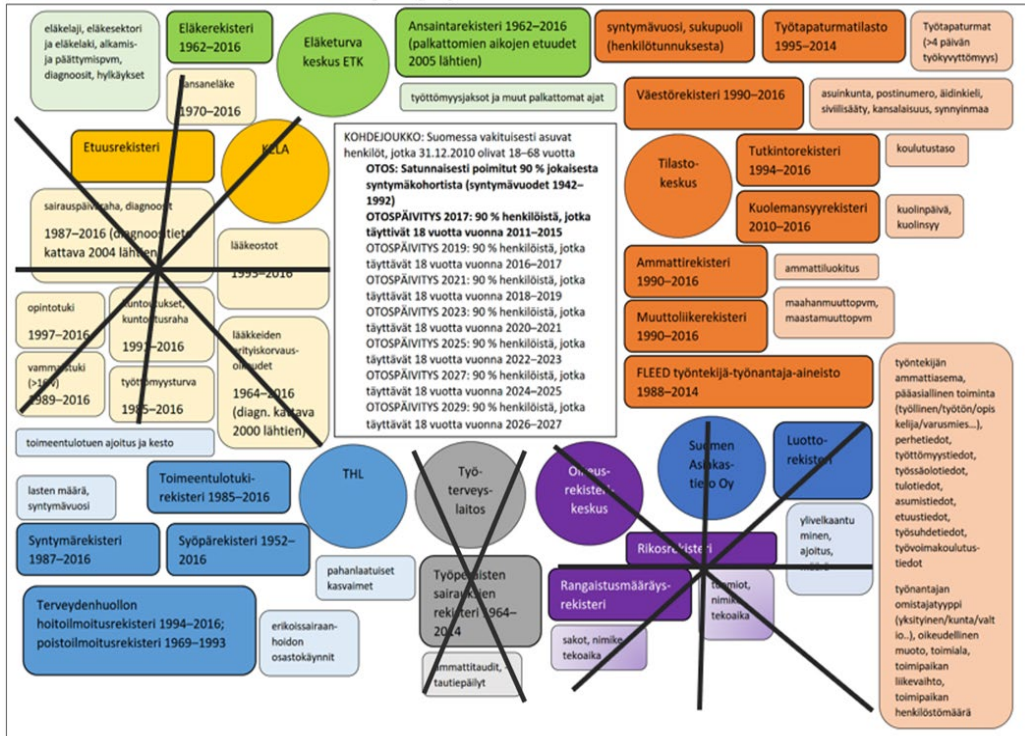
Tämän tutkimuksen tausta-aineistona toimi Työterveyslaitoksen käytössä oleva Työkyvyn tila Suomessa -aineisto. Tämä tutkimus on toteutettu käyttäen aineistoa Tilastokeskuksen etäkäyttöympäristössä (lupanumero TK-53-1198-17). Aineisto on pseudonymisoitu useiden rekisterien yhdistelmä, joka kattaa 90 prosentin satunnaisotoksen työikäisestä väestöstä. Tämän tutkimuksen tarpeisiin on hyödynnetty vuosien 2010 ja 2014 välisiä rekisteritietoja.

Aineiston perustana toimi Tilastokeskuksen Väestörekisteri, johon on liitetty muiden rekisterienpitäjien tietoja. Alkuperäisen aineiston muodostivat Tilastokeskuksen Väestörekisteri, Tutkintorekisteri, Kuolemansyyrekisteri, Ammattirekisteri, Muuttoliikerekisteri, FLEED työntekijä-työntekijä-aineisto, Työtapaturomatilasto sekä Eläketurvakeskuksen Ansaintarekisteri ja Eläkerekisteri. Lisäksi näihin on yhdistetty Työterveyslaitoksen ylläpitämä Työperäisten sairauksien rekisteri sekä Terveyden ja hyvinvoinnin laitoksen hallinnoima Toimeentulotukirekisteri ja Terveydenhuollon hoitoilmoitusrekisteri. Tiedot on yhdistetty yksilötasolla käyttäen pseudonymisoituja henkilötunnusteita. Lopputuloksena on laaja rekisteriaineisto, jonka avulla voidaan tarkastella työkyvyn tilannetta ja kehitystä suomalaisessa työikäisessä väestössä.

### 2.2 Tutkimuksessa hyödynnetty aineisto

Aineistosta koottiin tämän tutkimuksen tarpeisiin soveltuva rekisteriseuranta työikäiselle väestölle vuosina 2010–2014 (Kuva 1). Työikäinen väestö on Tilastokeskuksen määritelmän mukaisesti 15–74-vuotiaat, mutta tässä tarkasteltava väestö on rajattu 19–70-vuotiaisiin. Iäkkäämpien ryhmien tarkastelua perustelee se, että keskustelu eläkeiän nostosta on jatkuvaa. Koko tarkasteltavassa aineistossa oli yhteensä 3 167 732 henkilöä. Työkyvyttömyyteen liittyviä etuuksia saaneen väestön tarkastelussa oli 625 722 henkilöä.





Kuva 1. Työkyvyn tila Suomessa-rekisteriaineisto. Yliiviivattuja rekisteritietoja ei käytetty tässä tutkimuksessa.

Aineistosta muodostettiin 90 prosentin väestötöksen sisältävä, vuodet 2010–2014 kattava rekisteriseuranta liittyen työssäoloon, saatuihin etuuksiin tai erikoissairaanhoidon käyttöön. Muuttujia hyödynnettiin sekä jatkuvina muuttujina päivien tai kuukausien määrälle että luokittelevina muuttujina.

Terveysteen ja sairaanhoitoon liittyviä käytettyjä tietoja olivat hoitoilmoituspäivien määrä eli päivät, joina on saanut erikoissairaanhoidon. Työkyvyttömyyteen ja eläkkeisiin liittyviä etuustietoja olivat sairaus- ja osasairauspäiväraha, Kelan kuntoutusraha, täysi työkyvyttömyyseläke, kuntoutustuki eli määräaikainen työkyvyttömyyseläke, osatyökyvyttömyyseläke tai osakuntoutustuki eli määräaikainen osatyökyvyttömyyseläke. Lisäksi käytössä oli edellisiin liittyviä verrattain harvinaisempia etuuksia, kuten kuntoutusraha, osakuntoutusavustus, osakuntoutusraha, kuntoutusavustus ja kuntoutuskorotus. Kaikista edellä mainituista muuttujista tarkasteltiin ”päivien määrää per henkilö vuosina 2010–2014”.

Työssäolokuukausien ohella tarkasteltiin työttömyyskuukausia sekä kuukausia, jolloin on saanut toimeentulotukea. Toimeentulotuki myönnettiin vuosina 2010–2014 kunnan sosiaalitoimesta, joten kyseisen muuttuja indikoi myös aikuissosiaalityön asiakkuutta. Näiden lisäksi tutkimuksessa on käytetty tietoa perhe-etuuksista, kuten vanhempainrahojen ja sairauspäivärahan aikaisen vanhempainrahan käytöstä. Nämä tiedot koostettiin muotoon ”päivää per henkilö vuosina 2010–2014”.

Luokittelumuuttujina hyödynnettiin useita taustatietoja, kuten sukupuoli, asuntokunnan koko (1, 2 tai 3, enemmän), ikäluokat (alle 24-, 24-34-, 35-44-, 45-54-, 55-64- ja yli 65-vuotiaat, ikä on raportoitu myös jatkuvana muuttujana), alle 18-vuotiaiden lasten määrä perheessä (ei lasta tai ei tietoa, 1 lapsi, 2-3 lasta, 4 tai enemmän lasta), koulutusaste perustuen korkeimpaan tai viimeksi suoritettuun tutkintoon (toinen aste, erikoisammattitutkinto, alin korkea-aste, AMK- tai kandidaatin tutkinto, ylempi AMK- tai maisterintutkinto, lisensiaatti tai tohtorin tutkinto) tai koulutusasteen puuttumiseen, joka tulkittiin koulutuksen puutteeksi, kuntaryhmä (kaupunki, taajama, maaseutu), kansallisuus (Suomi, EU tai ei-EU tai ei tietoa), äidinkieli (suomi, ruotsi tai muu). Lisäksi Suomeen ulkomailta muuttaneille henkilöille on laskettu maassaoloaika (suomalainen tai ei tietoa maahantulosta, 0–4 vuotta maassa, 5–7 vuotta maassa, 8–10 vuotta maassa, yli 10 vuotta maassa). Edellä mainitut tiedot koskevat vuotta 2014.

Vuoden pisimmästä pääasiallisesta toiminnasta muodostettiin kolme uutta muuttujaa, jotka ilmaisevat ovatko henkilöt päätoimisesti opiskelijoita, eläkeläisiä tai varusmies- tai siviilipalveluksessa. Tällöin muuttujista nähdään kaksiarvoisen jaottelun perusteella, onko henkilö päätoimeltaan opiskelija, eläkeläinen tai ase- tai siviilipalveluksessa. Esimerkiksi päätoimisesti työssä olevilla henkilöillä muuttujat saavat arvon nolla. Tieto koskee vuotta 2014.

### 3 Menetelmät

Ryhmittelymenetelmien tavoitteena on jaotella havaintoyksiköt klustereihin eli ryhmiin määriteltyjen muuttujien samankaltaisuuden perusteella. Tässä tutkimuksessa havaintoyksikköinä toimivat yksittäiset henkilöt. Menetelmiä hyödynnetään usein suurien aineistojen kanssa, sillä niiden avulla voidaan jalostaa tietoa havainnollisempaan muotoon. Tällöin toisiaan vastaavat havainnot on ryhmitelty omiksi ryhmikseen. Tässä tutkimuksessa hyödynnetyssä käyttötarkoituksessa ryhmittely mahdollistaa aineistolähtöisesti luotujen ryhmien luomisen, ryhmien kuvailun sekä täten ryhmien välisten erojen tutkimisen.

K-means-menetelmä on niin sanottu kova ryhmittelymenetelmä, jossa yksi havainto voi sisältyä vain yhteen ryhmään. K-means-klusteroinnin avulla toteutettu ryhmittely perustuu havaintojen välisiin etäisyyksiin ja ryhmien keskipisteisiin, jotka perustuvat keskiarvoon. Jokaisen havainnon etäisyys valitusta ryhmien keskipisteestä minimoidaan. Toisin sanoen ryhmien keskiarvoa päivitetään niin kauan, että havaintojen etäisyydet ryhmän keskiarvosta ovat minimoidut ja jokainen havainto kuuluu ryhmään, jonka keskiarvoon perustuva keskipiste on etäisyyden perusteella mahdollisimman lähellä. Näin yksikään havainto ei jää ryhmien ulkopuolelle.

K-prototypes-ryhmittely kykenee ottamaan huomioon myös kategoriset muuttujat. Menetelmässä kategorisien muuttujien havaintojen välinen samankaltaisuus todetaan vertailemalla parittaisten havaintojen luokkia, jolloin samassa luokassa olevat parit saavat arvon 0 ja eri luokassa olevat arvon 1 (Huang, 1998). Laskemalla parittaiset vertailut yhteen pienempi arvo viittaa samankaltaisuuteen ja suurempi arvo erilaisuuteen. Tällöin samaa luokkaa edustavat havaintoyksiköt sijoittuvat suuremmalla todennäköisyydellä samoihin ryhmiin. Yhdistämällä kategoristen muuttujien samankaltaisuus tavallisessa K-means-laskennassa käytettävään Euklidiseen etäisyyteen, kerrotaan samankaltaisuus painokertoimella, joka estimoidaan jatkuvien muuttujien arvojen perusteella. Tätä yhteistä etäisyyttä pyritään minimoimaan iteroiden siirtämällä ryhmien keskiarvoon perustuvia keskustoja.

Ryhmien määrä, joihin havainnot jaotellaan, tulee valita etukäteen. Validointia tehdään laskemalla klustereiden sisäisiä etäisyyksiä klusterin keskipisteestä (Halkidi ym., 2001). Klustereiden havaintojen sisäisten etäisyyksien summa tulisi olla mahdollisimman pieni, jolloin klustereiden hajonta pienenee ja klusterit ovat tiiviimpiä. Tämän tutkimuksen osalta ryhmien määrän validointia voitiin tehdä hyödyntäen yhden prosentin satunnaisotosta kokonaisaineistosta.

## 4 Tulokset

### 4.1 Taustamuuttajat

Tutkimusaineistomme toimi 90 prosentin satunnaisotos työikäisestä väestöstä. Otos koostui 3 167 732 henkilöstä.

Henkilöiden keskimääräinen ikä oli hieman yli 45 vuotta (Taulukko 2). Puolet väestöstä oli naisia, noin 70 prosenttia asui kaupungissa. 2,7 prosentilla oli ulkomaalaisstatus tai puuttuva maatieto. Muita kuin suomen- tai ruotsinkielisiä oli 8,2 prosenttia. Viidesosa henkilöistä oli 55–64-vuotiaita. Alle 18-vuotiaita lapsia ei ollut noin 70 prosentilla. Viisi prosenttia oli opiskelijoita, 18 prosenttia eläkeläisiä. Hieman alle puolet omasi toisen asteen koulutuksen. Vähintään alimman korkea-asteen koulutus oli noin kolmasosalla väestöstä. (Taulukko 1)

Työttömyyttä oli kertynyt keskimäärin 4,4 kuukautta vuosina 2010–2014. Työssäoloa oli kertynyt keskimäärin 39,2 kuukautta. Toimeentulotukea oli saatu alle kaksi kuukautta. (Taulukko 2).

Sairauspäivärahopäiviä oli kertynyt keskimäärin yli 12 päivää vuosina 2010–2014, perhe-etuuksia lähes 50 päivää. Lisäksi hoitoilmoituspäiviä oli kertynyt keskimäärin 10. (Taulukko 2). Monet taulukossa esitetyt etuudet ovat erittäin harvinaisia koko työikäisessä väestössä, jolloin niitä kuvaavat "päivää per henkilö"-keskiarvot ovat nolla tai lähellä nollaa.

<b>MUUTTUJA</b>	<b>PROSENTTIA</b>
<b>Sukupuoli:</b> nainen	50,0
<b>Kuntaryhmä:</b> kaupunki	70,2
<b>Kansalaisuus:</b> ulkomaalaiset tai ei tietoa	2,7
<b>Maassaoloaika:</b> 0–4 vuotta maassa	0,2
<b>Kieli:</b> muu	8,2
<b>Iät jaoteltuna:</b> 55–64 -vuotiaat	20,8
<b>Lasten lukumäärä:</b> ei lapsia/ei tietoa	70,2
<b>Asuntokunnan koko:</b> Asuntokunta 3 tai enemmän	39,3
<b>Opiskelija:</b> on opiskelija	5,2
<b>Eläkeläinen:</b> on eläkeläinen	18,3
<b>Armeija tai siviilipalvelus:</b> on armeijassa/siviilipalveluksessa	0,2
<b>Korkein tai viimeisin koulutus:</b> toinen aste	46,7
<b>alin korkea-aste</b>	11,2
<b>AMK- tai kandidaatin tutkinto</b>	12,1
<b>ylempi AMK- tai maisterin tutkinto</b>	10,2

Taulukko 1. 90 % satunnaisotos työikäisestä väestöstä, N=3 167 732. Kategoriset muuttujat, osuudet vuonna 2014.

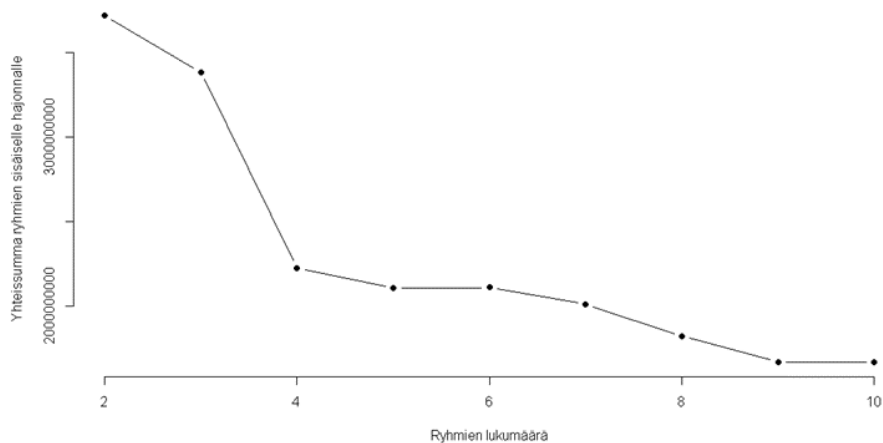
<b>MUUTTUJAT</b>	<b>KESKIARVO (HAJONTA)</b>
<b>Työttömyys, kuukaudet</b>	4,4 (10,4)
<b>Työssäolo, kuukaudet</b>	39,2 (23,9)
<b>Toimeentulotuki, kuukaudet</b>	1,8 (7,7)
<b>Kelan kuntoutusraha, päivät</b>	0,9 (18,4)
<b>Sairauspäiväraha, päivät</b>	12,4 (48,9)
<b>Osasairauspäiväraha, päivät</b>	0,5 (6,0)
<b>Perhe-etuudet, päivät</b>	47,5 (188,3)
<b>Täysi työkyvyttömyyseläke, päivät</b>	33,0 (206,6)
<b>Kuntoutustuki, päivät</b>	4,7 (69,3)
<b>Osatyökyvyttömyyseläke, päivät</b>	3,0 (61,0)
<b>Osakuntoutustuki, päivät</b>	0,7 (24,7)
<b>Kuntoutusraha työssä olevalle, päivät</b>	2,0 (35,5)
<b>Osakuntoutusavustus, päivät</b>	0,0 (0,1)
<b>Osakuntoutusraha, päivät</b>	0,0 (4,0)
<b>Kuntoutusavustus, päivät</b>	0,0 (1,8)
<b>Kuntoutuskorotus, päivät</b>	0,4 (12,7)
<b>Hoitoilmoitukset, päivät</b>	10,1 (32,8)
<b>Ikä, vuonna 2014</b>	45,2 (15,0)

Taulukko 2. 90 % satunnaisotos työkäisestä väestöstä, N=3 167 732. Jatkuvat muuttajat, keskiarvo ja hajonta vuosille 2010-2014.

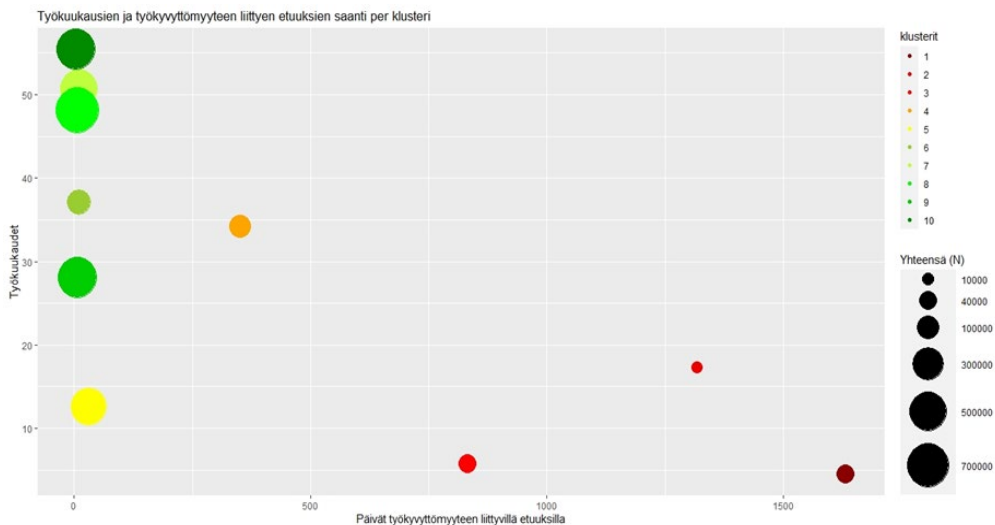
## 4.2 Ryhmittely koko väestölle

K-means-ryhmittely toteutettiin ensin koko väestölle. Sopivaksi ryhmien lukumääräksi arvioitiin kymmenen klusteria. Arvio perustui validointiin, joka tehtiin hyödyntäen yhden prosentin satunnaisotosta koko väestöstä. Validoinnissa havaittiin kymmenen ryhmän ratkaisun olevan sopivin mahdollinen klustereiden lukumäärä, koska tällä määrällä klusterien sisäinen hajonta oli lähimpänä nolaa (Kuva 2). Hajonnan pienenemisen perusteella voidaan olettaa ryhmien määrän kuvaavan mahdollisimman hyvin ryhmittelyssä käytettyjen muuttujien jakaumia. Toisiaan muistuttavat havainnot muodostuivat omiksi ryhmikseen.

K-Means-ryhmittelyn tuloksena muodostuneet ryhmät on käsitelty tässä pääpiirteittäin, ryhmä kerrallaan. Jatkuvista muuttujista on esitetty ryhmäkohtaisia keskiarvoja ja hajontalukuja. Kategorisista muuttujista on esitetty ryhmää kuvaavia prosenttiosuuksia.



Kuva 2. Koko 90 % väestöotosta koskevien klusterointiratkaisujen raja-arvoja. Y-akselilla on klusterien sisäisen hajonnan summa, kun ryhmien määrää muutetaan kahden ja kymmenen klusterin välillä. X-akseli kuvaa ryhmien määrää.



Kuva 3. 90 % väestöotus. Klusteriratkaisun ryhmät ja niiden sisältämät henkilömäärä työkuukausien ja työkyvyttömyysetuuspäivien mukaan. Y-akselilla Työkuukaudet vuosina 2010-2014, X-akselilla päivät työkyvyttömyysetuuskilla vuosina 2010-2014.

Ensimmäinen klusteri, "työkyvyttömät eläkeläiset" muodostui noin 41 000 henkilöstä, joista hieman alle puolet oli naisia. Klusteriin kuuluvien ikä oli keskimäärin 64 vuotta. Henkilöistä 96 prosenttia oli eläkkeellä. 60 prosenttia henkilöistä asui kaupungissa. Vain yhdellä prosentilla heistä oli alle 18-vuotiaita lapsia. Ulkomaalaistaustaisten osuus oli 0,3 prosenttia. Toisen asteen koulutus oli noin 45 prosentilla, alemman korkea-asteen koulutus yhdeksällä, alempi tai ylempi korkeakoulututkinto hieman alle viidellä prosentilla (Taulukko 3). Puuttuva koulutustieto oli noin 41 prosentilla. Yksin asuvia oli noin 37 prosenttia. (Taulukko 11, liitteenä).

Klusteriin kuuluvilla henkilöillä oli keskimäärin neljä työssäolokuukautta vuosien 2010-2014 aikana. Työttömyyskuukausia oli vain vähän, noin 0,3 kappaletta. Toimeentulotukea oli saatu keskimäärin yhden kuukauden ajan. Perhe-etuuksien käyttö oli vähäistä, keskimäärin 0,3 päivää per henkilö. (Taulukko 4).

Työkyvyttömyyteen liittyvistä etuuksista klusteriin kuuluvilla henkilöillä oli ollut keskimäärin 1600 päivää työkyvyttömyyseläkettä vuosina 2010-2014, eli he olivat olleet työkyvyttömyyseläkkeellä miltei koko ajanjakson. Määräaikaisten tai osa-aikaisten työkyvyttömyyteen liittyvien tukien käyttö oli vähäistä. Sairauspäivärahaa oli saatu hieman alle 19 päivää. Hoitoilmoituspäiviä, eli erikoissairaanhoidon käyttöä, oli 31. (Taulukko 5).

Ryhmään kuuluvat henkilöt olivat keskimäärin jo eläkeiässä. Ryhmässä oli siirrytty eläkkeelle työkyvyttömyyseläkkeen kautta. Työkyvyttömyyseläkkeellä oli oltu keskimäärin miltei koko ajanjakso.

Ryhmä sisällytettiin mukaan työkyvyttömyyteen liittyviä etuuksia saaneiden ryhmittelytarkasteluun, joka on seuraavassa luvussa.



KLUSTERIT	1 (N=4 0845)	2 (N=93 22)	3 (N=427 73)	4 (N=895 26)	5 (N=443 256)	6 (N=132 924)	7 (N=509 841)	8 (N=764 606)	9 (N=561 590)	10 (N=573 049)
<b>MUUTTUJAT</b>										
<b>Naiset, %</b>	48,6	57,0	48,5	46,3	66,0	97,8	76,8	40,0	37,3	29,4
<b>Kaupunkilaiset, %</b>	60,4	70,3	60,8	62,7	65,8	69,4	67,3	72,1	76,1	71,0
<b>Ulkomaalaiset, %</b>	0,30	1,50	0,40	1,90	1,80	5,30	3,20	2,70	2,20	3,30
<b>0-4 vuotta Suomessa asuneet, %</b>	0,0	0,2	0,0	0,1	0,1	0,2	0,2	0,2	0,2	0,2
<b>Muun kuin suomen- tai ruotsinkieliset, %</b>	3,9	5,5	4,7	6,1	7,2	11,0	9,4	7,7	7,1	10,0
<b>55 -64-vuotiaat, %</b>	58,4	9,0	20,2	45,3	22,6	0,1	1,4	19,0	13,8	44,1
<b>Toisen asteen koulutus, %</b>	44,7	57,2	39,3	57,3	25,5	40,2	49,1	64,1	70,1	15,2
<b>Alimman korkea-asteen koulutus, %</b>	9,0	8,3	9,2	10,1	14,0	5,1	15,4	6,4	6,1	18,3
<b>AMK- tai kandidaatin tutkinto, %</b>	2,7	8,0	3,0	5,0	5,8	24,1	15,1	15,3	6,8	14,9
<b>Ylempi AMK- tai maisterin tutkinto, %</b>	2,1	4,4	2,4	2,9	4,9	16,2	9,7	4,6	1,5	31,7
<b>Eläkeläiset, %</b>	96,0	53,5	96,4	30,0	86,0	0,6	0,8	3,5	7,2	2,7
<b>Ei alle 18-vuotiaita lapsia, %</b>	99,0	72,3	99,1	83,4	99,0	1,0	8,9	91,3	94,5	60,3
<b>Opiskelijat, %</b>	0,2	13,6	0,2	3,9	1,3	7,6	3,7	4,2	16,3	0,5
<b>Ase- tai siviilipalveluksessa, %</b>	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,1	0,0
<b>Asuntokunnassa 3 tai enemmän, %</b>	8,3	31,9	8,4	24,8	7,6	94,8	94,3	13,3	18,2	64,4
<b>Ikä, vuonna 2014 (Keskihajonta)</b>	63,68 (3,2)	41,07 (9,93)	65,55 (3,97)	50,55 (10,95)	62,47 (10,37)	33,84 (5,65)	41,78 (8,74)	40,11 (13,63)	36,52 (15,83)	49,42 (10,32)

Taulukko 3. 90 % Väestöotos, N=3 167 732. Kymmenen ryhmän ratkaisu: ikä sekä taustamuuttujia, prosenttiosuus klusterissa. Sarakkeet edustavat klustereita, rivit kuvaavat klustereiden piirteitä eri taustatekijöittäin.

<b>KLUSTERI T MUUTTUJ AT</b>	<b>1 (N=40 845)</b>	<b>2 (N=9 322)</b>	<b>3 (N=42 773)</b>	<b>4 (N=89 526)</b>	<b>5 (N=44 3256)</b>	<b>6 (N=13 2924)</b>	<b>7 (N=50 9841)</b>	<b>8 (N=76 4606)</b>	<b>9 (N=56 1590)</b>	<b>10 (N=57 3049)</b>
<b>Työttömyys</b>	0,3 (2,6)	5,1 (8,2)	1,0 (4,6)	9,2 (13,4)	2,7 (8,9)	3,9 (7,2)	3,5 (8,6)	5,0 (10,7)	7,9 (14,0)	2,2 (7,7)
<b>Työssäolo</b>	4,5 (12,0)	17,4 (19,9)	5,7 (12,8)	34,2 (22,1)	12,6 (19,2)	37,0 (22,7)	50,6 (16,9)	48,1 (17,3)	28,1 (22,8)	55,4 (12,6)
<b>Toimeentulotuki</b>	1,0 (4,9)	7,1 (12,0)	1,2 (5,5)	5,0 (12,3)	1,2 (6,0)	2,6 (8,8)	1,1 (5,6)	1,1 (5,4)	4,9 (12,7)	0,4 (3,6)
<b>Perhe-etuudet</b>	0,3 (14,6)	24,8 (122)	0,2 (11,2)	3,3 (27,0)	0,1 (3,8)	851,2 (317,4)	62,1 (118,8)	1,9 (11,4)	0,5 (7,2)	5,9 (21,1)

Taulukko 4. 90 % väestöotos. Kymmenen ryhmän ratkaisu: Työttömyys-, työssäolo- ja toimeentulotukikukaudet sekä päivät perhe-etuuksilla. Sarakkeet edustavat klustereita, rivit kuvaavat klustereiden piirteitä eri taustatekijöittäin.

KLUSTERIT MUUTTUJAT	1 (N=4 0845)	2 (N=9 322)	3 (N=4 2773)	4 (N=8 9526)	5 (N=44 3256)	6 (N=13 2924)	7 (N=50 9841)	8 (N=76 4606)	9 (N=56 1590)	10 (N=57 3049)
Työkyvyttömyyseläke	1613,0 (212)	0,1 (7,0)	793,3 (233)	8,0 (47,3)	8,1 (46,4)	0,1 (7,2)	0,0 (3,1)	0,1 (4,3)	0,5 (11,3)	0,1 (3,9)
Kuntoutustuki	0,2 (9,8)	1143 (415)	0,2 (7,2)	33,8 (103)	0,4 (10,9)	1,2 (23,9)	0,5 (13,0)	0,3 (10,5)	0,6 (13,8)	0,2 (7,1)
Osatyökyvyttömyyseläke	0,3 (19,7)	0,0 (0,0)	0,2 (9,9)	5,0 (75,6)	19,55 (155,7)	0,0 (2,2)	0,0 (5,7)	0,2 (11,0)	0,2 (11,7)	0,4 (17,4)
Osakuntoutustuki	0,0 (2,9)	1,1 (31,4)	0,1 (5,8)	10,6 (94,2)	0,2 (14,9)	0,3 (16,7)	0,8 (28,4)	0,4 (17,7)	0,4 (19,7)	0,2 (14,1)
Kuntoutusraha	0,3 (6,7)	22,8 (111)	0,3 (8,7)	44,3 (176,7)	0,2 (6,6)	0,7 (21,4)	1,5 (28,4)	0,7 (15,8)	0,6 (16,1)	0,4 (11,5)
Osakuntoutusavustus	0,0 (0,0)	0,0 (0,3)	0,0 (0,0)	0,0 (0,0)	0,0 (0,1)	0,0 (0,0)	0,0 (0,0)	0,0 (0,0)	0,0 (0,3)	0,0 (0,0)
Osakuntoutusraha	0,0 (0,0)	0,2 (9,2)	0,0 (0,0)	0,1 (9,6)	0,0 (0,0)	0,0 (3,5)	0,1 (6,5)	0,0 (3,4)	0,0 (2,4)	0,0 (3,5)
Kuntoutusavustus	0,0 (0,0)	0,2 (8,4)	0,0 (0,0)	0,1 (7,6)	0,0 (1,6)	0,0 (0,0)	0,0 (1,8)	0,0 (0,4)	0,0 (0,8)	0,0 (1,7)
Kuntoutuskorotus	0,2 (6,1)	47,2 (173)	0,1 (5,3)	3,6 (29,0)	0,2 (7,5)	0,1 (4,5)	0,2 (8,2)	0,1 (6,3)	0,1 (7,0)	0,1 (5,0)
Kelan kuntoutusraha	0,1 (2,4)	5,2 (40,7)	0,2 (4,6)	6,1 (45,1)	0,3 (8,4)	0,5 (13,4)	0,8 (16,1)	0,7 (17,1)	1,5 (27,5)	0,4 (7,4)
Hoitoilmoitukset	30,9 (66,6)	110,4 (134)	29,9 (61,9)	45,4 (80,3)	17,4 (50,1)	10,0 (19,0)	6,8 (16,1)	5,4 (14,5)	8,8 (29,1)	4,8 (12,1)
Sairauspäiväraha	18,6 (61,2)	145,0 (131)	38,2 (88,8)	243,2 (87,0)	3,2 (17,4)	7,9 (31,8)	6,5 (23,9)	5,1 (17,7)	4,4 (18,0)	3,0 (14,1)
Osasairauspäiväraha	0,1 (2,4)	1,5 (9,7)	0,2 (3,1)	3,8 (16,3)	0,2 (3,2)	0,3 (4,2)	0,7 (7,2)	0,5 (5,8)	0,3 (4,5)	0,4 (5,2)

Taulukko 5. 90 % väestöotos. Kymmenen ryhmän ratkaisu: työkyvyttömyyteen liittyvät etuudet, päivää etuuksilla vuosina 2010-2014. Sarakkeet edustavat klustereita, rivit kuvaavat klustereiden piirteitä eri taustatekijöittäin.

Toinen klusteri, "kuntoutuksessa olevat osatyökykyiset", koostui yhteensä 9322 henkilöstä, joista naisia oli 57 prosenttia. Ryhmään kuuluvien henkilöiden keskimääräinen ikä oli 41 vuotta. Eläkkeellä heistä oli 53,5 prosenttia. Ryhmään kuului myös opiskelijoita, jopa 14 prosenttia. 70 prosenttia henkilöistä asui kaupungissa. Kolmasosalla heistä oli alle 18-vuotiaita lapsia. Ulkomaalaistaustaisten osuus oli 1,5 prosenttia. Toisen asteen koulutus oli noin 57 prosentilla. Vähintään alimman korkeakoulututkinnon omasi 21 prosenttia. (Taulukko 3). Puuttuva koulutustieto oli noin 21 prosentilla. Yksin asuvia oli noin 39 prosenttia. (Taulukko 11, liitteenä).

Ryhmän henkilöillä oli vuosina 2010-2014 työssäolokuukausia keskimäärin noin 17. Työttömyyskuukausia oli kertynyt viisi. Toimeentulotukea oli saatu seitsemänä kuukautena. Perhe-etuuksia oli maksettu keskimäärin 25 päivää. (Taulukko 4).

Työkyvyttömyyteen liittyvistä etuuksista klusteriin kuuluvilla henkilöillä oli ollut keskimäärin 1143 kuntoutustukipäivää, 23 päivää kuntoutusrahaa, 47 päivää kuntoutuskorotusta. Klusteriin kuuluvilla henkilöillä oli myös keskimäärin 145 päivää sairauspäivärahopäivää ja 1,5 osasairauspäivärahopäivää vuosina 2010-2014. Täydet työkyvyttömyyseläkkeet olivat harvinaisia, keskimäärin 0,1 päivää henkilöä kohden. Heillä oli verrattain paljon erikoissairaanhoidossa hoidettavia terveysongelmia, eli noin 110 hoitoilmoituspäiviä. (Taulukko 5).

Toisessa ryhmässä työkyvyttömyyteen liittyvien etuuksien käyttö oli erityisesti määrätai osa-aikaisten työkyvyttömyysetuuksien käyttöä. Hoitoilmoituspäiviä oli paljon verrattuna muihin ryhmiin. Ryhmässä oli terveyden ongelmia, jotka ilmenivät hoitoilmoitusten lisäksi pitkinä sairauslomina ja kuntoutuksen tarpeena. Täydet työkyvyttömyyseläkkeet olivat kuitenkin harvinaisia. Ryhmä on sisällytetty mukaan toiseen ryhmittelytarkasteluun.

Ensimmäisen ryhmän tapaan kolmatta ryhmää voi kuvailla nimellä "työkyvyttömät eläkeläiset". Ryhmän 42 773 henkilöstä 49 prosenttia on naisia. Ryhmään kuuluvien henkilöiden keski-ikä oli 66 vuotta. Henkilöistä 96 prosenttia oli eläkeläisiä ja 61 prosenttia asui kaupungissa. Vain hieman alle prosentilla oli alle 18-vuotiaita lapsia. Toisen asteen koulutus oli 39 prosentilla, vähintään alimman korkea-asteen koulutus noin 15 prosentilla (Taulukko 3). Puuttuva koulutustieto oli noin 46 prosentilla. Yksin asuvien osuus oli 34 prosenttia. (Taulukko 11, liitteenä).

Vuosina 2010-2014 työssäolokuukausia oli noin kuusi ja työttömyyskuukausia yksi. Toimeentulotukikuukausia oli keskimäärin yksi. Perhe-etuuksien käyttö oli harvinaista, vain 0,2 päivää henkilöä kohden. (Taulukko 4).

Ryhmään kuuluvilla henkilöillä oli kertynyt keskimäärin alle 800 työkyvyttömyyseläkepäivää. Sairauspäivärahopäiviä oli kertynyt 38. Ryhmään kuuluvilla henkilöillä oli vain vähän määräaikaisten työkyvyttömyyteen liittyvien etuuksien käyttöä. (Taulukko 5).

Merkittävin ero ensimmäiseen ryhmään verrattuna löytyy työkyvyttömyyteen liittyvistä etuuksista. Ensimmäisessä ryhmässä oli kertynyt keskimäärin kaksinkertainen pituus täydellä työkyvyttömyyseläkkeellä oloa verrattuna kolmanteen klusteriin. Henkilöillä on

keskimäärin alle 800 työkyvyttömyyseläkepäivää, kun ensimmäisessä ryhmässä työkyvyttömyyseläkepäiviä oli kaksinkertainen määrä, yli 1600.

Työkyvyttömyysetuuksien käytön perusteella henkilöiden työkyvyttömyys oli voinut ilmetä ensimmäisen ryhmän henkilöitä myöhemmin. Ero työkyvyttömyyseläkkeen keskimääräisessä pituudessa voi selittyä myös sillä, että ryhmään kuuluvat henkilöt olivat keskimäärin kaksi vuotta vanhempia kuin ensimmäisen ryhmän henkilöt, jolloin työkyvyttömyyseläke on ehtinyt loppua ja varsinainen eläke alkaa jo aiemmin aikavälillä 2010-2014. Kolmas ryhmä on otettu mukaan toiseen ryhmittelytarkasteluun.

Neljäs ryhmä, "kuntoutuksessa tai sairauspäivärahalla olevat osatyökykyiset", koostui 89 526 henkilöstä, joista 46 prosenttia oli naisia. Ryhmään kuuluvien henkilöiden keskimääräinen ikä oli 51 vuotta. Henkilöistä 30 prosenttia oli eläkkeellä, 4 prosenttia opiskelijoina. Kaupungissa asui lähes 63 prosenttia henkilöistä. Alle 18-vuotiaita lapsia oli 17 prosentilla henkilöistä. Toisen asteen koulutus oli 57 prosentilla, vähintään alimman korkea-asteen koulutus oli 18 prosentilla. (Taulukko 3). Puuttuva koulutustieto oli noin 24 prosentilla. Yksin asuvia oli noin 29 prosenttia. (Taulukko 11, liitteenä).

Työssäolokuukausia henkilöillä oli keskimäärin noin 34 kuukautta, työttömyyskuukausia yhdeksän (Taulukko 4).

Täydellä työkyvyttömyyseläkkeellä oli oltu keskimäärin kahdeksan päivää henkilöä kohden. Kuntoutustukea oli saatu noin 34 päivää, osatyökyvyttömyyseläkettä viisi päivää ja määräraikaista kuntoutustukea noin 11 päivää. Kuntoutusrahopäiviä työelämästä tulevalle oli keskimäärin 44, Kelan kuntoutusrahopäiviä noin kuusi ja sairauspäivärahopäiviä noin 243 päivää. Osasairauspäivärahaa oli saatu keskimäärin neljä päivää. Hoitoilmoituspäiviä oli 45, joten erikoissairaanhoidon käyttö oli verrattain korkealla tasolla. (Taulukko 5).

Neljännessä ryhmässä työkyvyttömyyden haasteet eivät olleet eskaloituneet vielä täydeksi työkyvyttömyydeksi kaikilla henkilöillä. Erityisesti sairauspäivärahopäiviä oli kertynyt paljon. Myös hoitoilmoituspäiviä oli enemmän kuin useissa muissa klustereissa. Ryhmä otettiin mukaan toiseen ryhmittelytarkasteluun.

Viides ryhmä, "osatyökykyiset ikääntyvät", muodostui 443 256 henkilöstä, joista 66 prosenttia oli naisia. Ryhmään kuuluvien keskimääräinen ikä oli noin 62 vuotta. Henkilöistä 86 prosenttia oli eläkkeellä. 66 prosenttia asui kaupungissa. Alle 18-vuotiaita lapsia oli noin yhdellä prosentilla henkilöistä. Toisen asteen koulutus oli 26

prosentilla, vähintään alimman korkea-asteen koulutus 25 prosentilla. (Taulukko 3). Puuttuva koulutustieto oli noin 49 prosentilla. Yksin asuvia oli noin 25 prosenttia. (Taulukko 11, liitteenä).

Työssäoloa vuosina 2010-2014 oli kertynyt keskimäärin noin 13 kuukautta. Työttömyyskuukausia oli kertynyt noin kolme. Toimeentulotukikuukausia oli keskimäärin yksi. Perhe-etuuksia oli käytetty keskimäärin 0,1 päivää henkilöä kohden. (Taulukko 4).

Täyttä työkyvyttömyyseläkettä oli saatu keskimäärin kahdeksan päivää henkilöä kohden. Ryhmään kuuluvilla henkilöillä oli kaikista ryhmistä suurin keskiarvo osatyökyvyttömyyseläkepäivien määrä, noin 19 päivää. Muut määräaikaiset työkyvyttömyyteen liittyvät etuudet olivat harvinaisia. Sairauspäivärahopäiviä oli kertynyt noin kolme päivää, osasairauspäivärahoja vain 0,2 päivää. Hoitoilmoituspäiviä oli noin 17. (Taulukko 5).

Ryhmässä oli työkyvyttömyyteen liittyvien etuuksien käyttöä, erityisesti osatyökyvyttömyyseläkepäiviä, joita oli enemmän kuin muissa ryhmissä. Ryhmä oli suuri. Siihen kuului lähes puoli miljoonaa henkilöä. Klusteriin kuuluvilla henkilöillä oli keskimäärin vain vähän sairauspäiväraha- tai osasairauspäivärahopäiviä. Ryhmässä on siis paljon työkyvyttömyysetuuksia saaneita, mutta ei yksinomaan heitä. Ryhmä otettiin mukaan toiseen ryhmittelytarkasteluun.

Kuudes ryhmä, "perhe-etuudella olleet", koostui 132 924 henkilöstä, joista 98 prosenttia oli naisia. Keskimääräinen ikä oli 34 vuotta, ja lähes kaikilla (99 %) oli alle 18-vuotiaita lapsia. Eläkkeellä heistä oli yksi prosentti, opiskelemassa kahdeksan prosenttia. Maahanmuuttajien osuus oli kaikista ryhmistä suurin, 5 prosenttia. Henkilöistä 69 prosenttia asui kaupungissa. Toisen asteen koulutus oli hieman yli 40 prosentilla, vähintään alimman korkea-asteen koulutus 45 prosentilla. (Taulukko 3). Puuttuva koulutustieto oli noin 13 prosentilla. Yksin asuvia oli vain puoli prosenttia. (Taulukko 11, liitteenä).

Henkilöillä oli keskimäärin 37 työssäolokuukautta vuosina 2010–2014. Työttömyyskuukausia oli neljä. Toimeentulotukikuukausia oli noin 2,6. Perhe-etuuksia oli käytetty eniten kaikista ryhmistä, noin 851 päivää. (Taulukko 4).

Työkyvyttömyyteen liittyvät etuudet olivat harvinaisia, vaikkakin sairauspäivärahopäiviä oli keskimäärin kahdeksan henkilöä kohden. Hoitoilmoituspäiviä oli kertynyt kymmenen. (Taulukko 5).

Ryhmä koostui paljolti tilapäisesti työelämän ulkopuolella eli esimerkiksi vanhempainvapaalla olleista naisista. Ryhmää ei sisällytetty jatkotarkasteluun.

Seitsemäs ryhmä, ”perheelliset, työkykyiset työssäkäyvät”, koostui 509 841 henkilöstä, joista naisia oli 77 prosenttia. Henkilöiden keskimääräinen ikä oli 42 vuotta. Henkilöistä yksi prosentti oli eläkkeellä. Opiskelijoita oli neljä prosenttia. 67 prosenttia asui kaupungissa. 91 prosentilla henkilöistä oli alle 18-vuotiaita lapsia. Toisen asteen koulutus oli 49 prosentilla, vähintään alimman korkea-asteen koulutus 25 prosentilla. (Taulukko 3). Puuttuva koulutustieto oli noin 8 prosentilla. Yksin asuvia oli vain 0,1 prosenttia. (Taulukko 11, liitteenä).

Henkilöillä oli keskimäärin 51 työssäolokuukautta vuosina 2010–2014, työttömyyskuukausia oli neljä. Toimeentulotukikuukausia oli yksi. Perhe-etuksia oli käytetty keskimäärin 62 päivää. (Taulukko 4).

Henkilöillä oli vain vähän työkyvyttömyyteen liittyviä etuuksia. Sairauspäivärahopäiviä oli noin kuusi, osasairauspäivärahopäiviä yksi, hoitoilmoituspäiviä seitsemän. (Taulukko 5).

Ryhmää ei sisällytetty jatkotarkasteluun.

Kahdeksas ryhmä, ”lapsettomat työssäkäyvät työkykyiset”, koostui 764 606 henkilöstä, joista naisia oli 40 prosenttia. Henkilöiden keskimääräinen ikä oli 40 vuotta. Heistä 72 prosenttia asui kaupungissa. Eläkkeellä oli neljä prosenttia ja opiskelijana neljä prosenttia. Vain noin yhdeksällä prosentilla oli alle 18-vuotiaita lapsia. Toisen asteen koulutus oli 64 prosentilla, vähintään alimman korkea-asteen koulutus 26 prosentilla. (Taulukko 3). Puuttuva koulutustieto oli noin 8 prosentilla. Yksin asuvia oli noin 11 prosenttia. (Taulukko 11, liitteenä).

Työssäolokuukausia oli keskimäärin noin 48 ja työttömyyskuukausia keskimäärin viisi (Taulukko 4).

Työkyvyttömyyteen liittyvien etuuksien määrät olivat henkilöillä vähäisiä. Sairauspäivärahopäiviä oli kertynyt keskimäärin viisi. Myös hoitoilmoituspäiviä oli keskimäärin viisi. (Taulukko 5.)

Ryhmää ei sisällytetty jatkotarkasteluun

Yhdeksänteen ryhmään, "työhön kiinnittyvät nuoret", kuului 561 590 henkilöä, joista 37 prosenttia oli naisia. Henkilöiden keskimääräinen ikä oli 37 vuotta. Eläkkeellä henkilöistä oli seitsemän prosenttia, opiskelemissa 16 prosenttia ja varusmies- tai siviilipalveluksessa yksi prosentti. 76 prosenttia henkilöistä asui kaupungissa. Viidellä prosentilla oli alle 18-vuotiaita lapsia. Noin 70 prosentilla oli toisen asteen koulutus, vähintään alimman korkea-asteen koulutuksen omaasi 15 prosenttia. (Taulukko 3). Puuttuva koulutustieto oli noin 15 prosentilla. Yksin asuvia oli noin 70 prosenttia. (Taulukko 11, liitteenä).

Työssäolokuukausia oli noin 28 kuukautta, työttömyyskuukausia kahdeksan. Toimeentulotukikuukausia oli noin viisi, joka oli eniten vain vähän työkyvyttömyyteen liittyviä etuuksia saaneiden klustereista. (Taulukko 4).

Työkyvyttömyyteen liittyvät etuuksien käyttö oli vähäistä tässä ryhmässä, joskin Kelan kuntoutusrahaa oli saatu keskimäärin puolitoista päivää. Sairauspäivärahopäiviä oli keskimäärin neljä ja hoitoilmoituspäiviä oli yhdeksän. (Taulukko 5). Ryhmässä oli yhdessä kuudennen ryhmän kanssa vähiten työkyvyttömyyteen liittyviä etuuksia ja hoitoilmoituspäiviä.

Ryhmässä korostui muun muassa yksinasuminen. Siksi mukana on verrattain suuri määrä eläkkeellä olevia. He ovat lähes yksinomaan eläköityneet ikänsä puolesta, ei työkyvyttömyyden vuoksi. Ryhmää ei sisällytetty jatkotarkasteluun.

Kymmenenteen ryhmään, "korkeakoulutetut työkykyiset", kuului 573 049 henkilöä, joista 29 prosenttia oli naisia. Henkilöiden keskimääräinen ikä oli 49 vuotta.. Henkilöistä kolme prosenttia oli eläkkeellä ja opiskelemissa puoli prosenttia. Kaupungissa asui 71 prosenttia henkilöistä. 40 prosentilla oli alle 18-vuotiaita lapsia. Heillä oli usein muita ryhmiä korkeampi koulutusaste, kun toisen asteen koulutuksen saaneiden osuus oli vain 15 prosenttia ja vähintään alimman korkea-asteen koulutus oli yli kahdella kolmasosalla klusteriin kuuluvista henkilöistä. (Taulukko 3). Puuttuva koulutustieto oli noin 16 prosentilla. Yksin asuvia oli noin 14 prosenttia. (Taulukko 11, liitteenä).

Työssäolokuukausia heillä oli keskimäärin 55 kuukautta, työttömyyskuukausia noin kaksi. Toimentulotukikuukaudet olivat harvinaisia, lähes nollassa. Perhe-etuuksia oli saatu keskimäärin 6 päivää henkilöä kohden. (Taulukko 4).

Työkyvyttömyyteen liittyvien etuuksien käyttö oli vähäistä. Sairauspäivärahopäiviä oli keskimäärin kolme ja hoitoilmoituspäiviä keskimäärin viisi (Taulukko 5).

Ryhmää ei ole sisällytetty jatkotarkasteluun.



### 4.3 Koko väestön ryhmittely ja työssäolokuukaudet

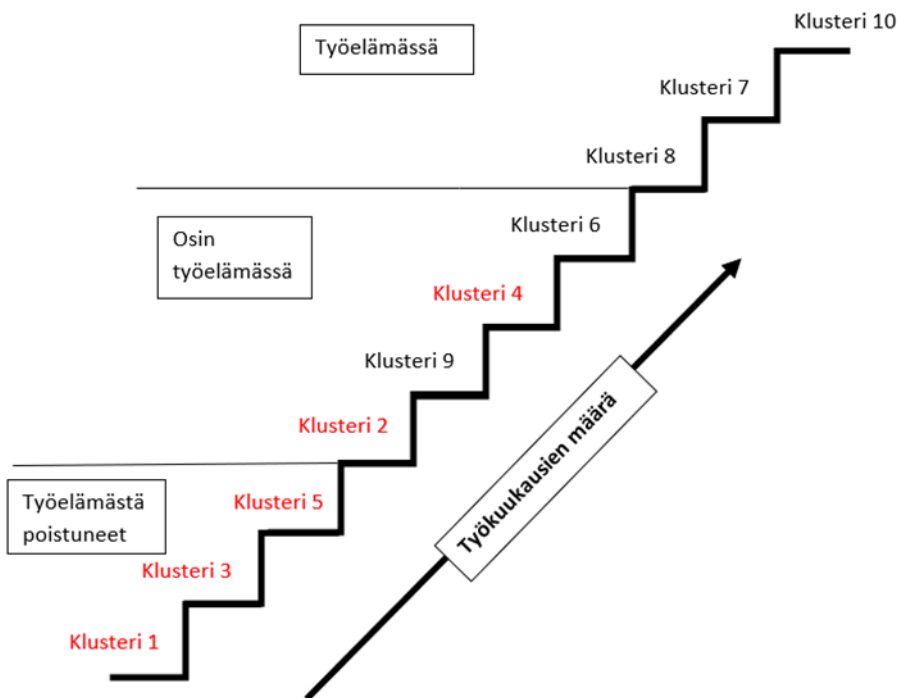
Kuvassa 3 on koko väestöä koskevan ryhmittelyratkaisun ryhmät järjestettynä vuosina 2010-2014 toteutuneiden työssäolokuukausien mukaisesti. Punaisella merkityt klusterit olivat niitä, joilla oli huomattavaa työkyvyttömyyteen liittyvien etuuksien saantia.

Ryhmät 1, 3 ja 5 olivat työssäolokuukausien mukaan tarkasteltuna työelämän ulkopuolella. Näihin ryhmiin kuului yhteensä 526 874 henkilöä. Heistä valtaosa (443 256 henkilöä) oli ryhmässä 5, jossa on työssäolokuukausia kertyi keskimäärin 12 kappaletta vuosina 2010-2014. Myös työkyvyttömyyteen liittyvien etuuksien käyttö oli vähäisempää ryhmässä 5. Keskimääräinen ikä kussakin ryhmässä oli yli 62 vuotta.

Ryhmät 2, 4, 6 ja 9 olivat osittain työelämässä. Ryhmissä 2 ja 4 oli yhteensä 98 848 henkilöä. Ryhmissä oli yleisesti saatu määräaikaista työkyvyttömyyteen liittyviä etuuksia. Ryhmässä 2 työssäolokuukausia kertyi noin 17, kun ryhmässä 4 niitä oli 34. Ryhmien 6 ja 9 keski-ikä oli lähes 35 vuotta. Ryhmässä 6 oli saatu verrattain paljon perhe-etuuksia. Ryhmässä 9 oli muita ryhmiä suurempi osuus päätoimisia opiskelijoita.

Ryhmien 7, 8 ja 10 henkilöt olivat työssäolokuukausien perusteella arvioituna työelämässä. Ryhmissä oli yhteensä 1 847 496 henkilöä, eli yli puolet koko työikäisestä väestöstä. Heillä oli vain vähän työkyvyttömyyteen liittyvien etuuksien saantia.

Ryhmät 1, 2, 3, 4 ja 5 on sisällytetty jatkotarkasteluun.



Kuva 4. 90 % väestötös, N=3 167 732. Kymmenen ryhmän ratkaisu. Ryhmät järjestetty vuosina 2010-2014 toteutuneiden työssäolokuukausien mukaisesti vähimmästä työssäolokuukausien määrästä enimpään.

#### 4.4 Ryhmittely työkyvyttömyyteen liittyviä etuuksia saaneiden väestölle

Koko aineiston ryhmittelyn jälkeen ryhmittelyä jatkettiin toistamalla K-means-ryhmittely vain työkyvyttömyyteen liittyviä etuuksia saaneiden ryhmille. Nämä olivat koko väestöä koskevan kymmenen klusterin ratkaisussa muodostetut ryhmät 1, 2, 3, 4 ja 5, joihin kuului yhteensä 625 722 henkilöä. Työkyvyttömyyteen liittyviä etuuksia saaneiden väestöä on kuvailtu taulukoissa 6 ja 7.

Työkyvyttömyyden etuuksia saaneet ovat koko 90 % väestötöstä useammin naisia, iäkkäämpiä, vailla koulutusta ja eläkeläisiä (Taulukko 6 ja taulukko 12, liitteenä). Työkyvyttömyyden etuuksia oli saatu hyvin laajalti. (Taulukko 7).

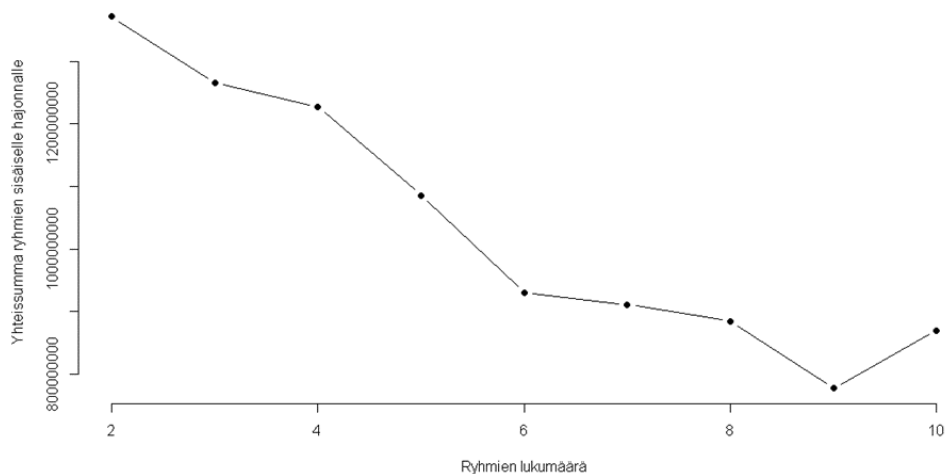
<b>MUUTTUJAT</b>	<b>PROSENTTIA</b>
<b>Sukupuoli:</b> nainen	60,7
<b>Kuntaryhmä:</b> kaupunki	64,7
<b>Kansalaisuus:</b> ulkomaalaiset tai ei tietoa	1,7
<b>Maassaoloaika:</b> 0–4 vuotta maassa	0,1
<b>Kieli:</b> muu	6,6
<b>Iät jaoteltuna:</b> 55–64-vuotiaat	27,8
<b>Lasten lukumäärä:</b> ei lapsia/ei tietoa	96,4
<b>Asuntokunnan koko:</b> Asuntokunta 3 tai enemmän	10,5
<b>Opiskelija:</b> on opiskelija	1,7
<b>Eläkeläinen:</b> on eläkeläinen	78,9
<b>Ase- tai siviilipalvelus:</b> on ase- tai siviilipalveluksessa	0,0
<b>Korkein tai viimeisin koulutus:</b> toinen aste	32,7
alin korkea-aste	12,7
<b>AMK- tai kandidaatin tutkinto</b>	5,3
ylempi AMK- tai maisterin tutkinto	4,2

Taulukko 6. Osatyökykyinen väestö, N=625 722. Kategoriset muuttujat, prosenttiosuudet vuonna 2014.

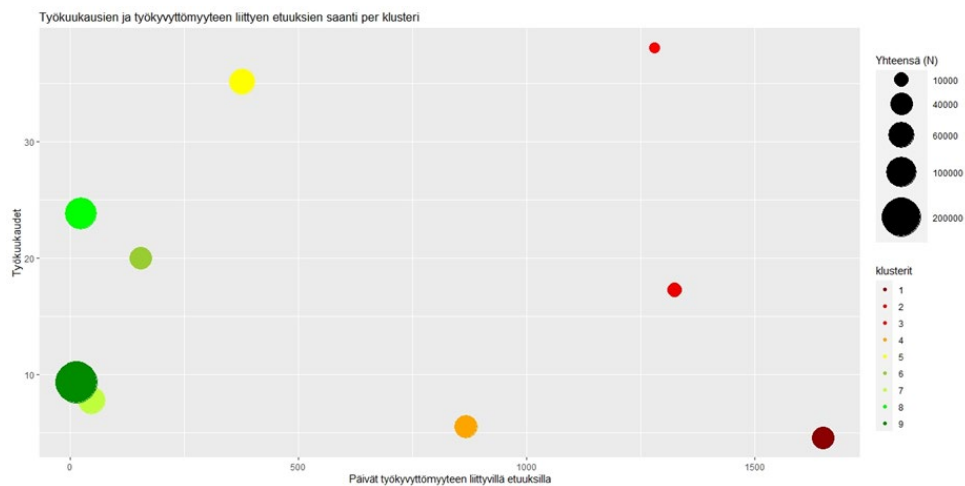
MUUTTUJAT	KESKIARVO (HAJONTA)
Työttömyys, kuukaudet	3,4 (9,6)
Työssäolo, kuukaudet	14,7 (20,7)
Toimeentulotuki, kuukaudet	1,8 (7,4)
Kelan kuntoutusraha, päivät	1,2 (19,3)
Sairauspäiväraha, päivät	43,0 (96,9)
Osasairauspäiväraha, päivät	0,7 (7,1)
Perhe-etuudet, päivät	0,9 (19,2)
Täysi työkyvyttömyyseläke, päivät	166,4 (440,2)
Kuntoutustuki, päivät	22,1 (152,8)
Osatyökyvyttömyyseläke, päivät	14,5 (134,4)
Osakuntoutustuki, päivät	1,7 (38,2)
Kuntoutusraha työssä olevalle, päivät	6,9 (70,2)
Osakuntoutusavustus, päivät	0 (0,1)
Osakuntoutusraha, päivät	0,0 (3,8)
Kuntoutusavustus, päivät	0,0 (3,3)
Kuntoutuskorotus, päivät	1,3 (25,4)
Hoitoilmoitukset, päivät	24,5 (61,1)
Ikä, vuonna 2014	60,7 (11,0)

Taulukko 7. Osatyökykyinen väestö, N=625 722. Jatkuvien muuttujien keskiarvo ja hajonta vuosille 2010-2014. Kuukautta, päivää tai vuosia.

Työkyvyttömyyteen liittyviä etuuksia saaneiden väestön ryhmittely päädyttiin tekemään yhdeksään ryhmään, jolloin ryhmien sisäisen hajonnan havaittiin lähenevän nolaa (Kuva 5). Yhdeksää ryhmää on kuvailtu seuraavaksi ryhmä kerrallaan. Jatkuvista muuttujista on esitetty ryhmäkohtaisia keskiarvoja ja hajontalukuja. Kategorisista muuttujista on esitetty ryhmää kuvaavia prosenttiosuuksia.



Kuva 5. Työkyvyttömyyteen liittyviä etuuksia saaneen väestön ryhmittelyratkaisujen raja-arvoja. Y-akselilla on ryhmien sisäisen hajonnan summa, kun ryhmien määrää muutetaan kahden ja kymmenen ryhmän välillä.



Kuva 6. Työkyvyttömyyteen liittyviä etuuksia saaneiden väestö. Ryhmittelyratkaisun ryhmät ja niiden sisältämät henkilömäärät järjestettynä työkuukausien ja työkyvyttömyysetuuspäivien mukaan. Y-akselilla Työssäolokuukaudet vuosina 2010-2014, X-akselilla päivät työkyvyttömyysetuuksilla vuosina 2010-2014.

Ensimmäinen ryhmä, ”työkyvyttömät eläkeläiset” muodostui noin 40 000 henkilöstä, joista hieman alle puolet oli naisia. Ryhmään kuuluvien ikä oli keskimäärin 64 vuotta. Henkilöistä 96 prosenttia oli eläkkeellä. 61 prosenttia asui kaupungissa. Vain yhdellä prosentilla oli alle 18-vuotiaita lapsia. 46 prosentilla oli korkeintaan toisen asteen koulutus. Vähintään alimman korkea-asteen koulutus oli 15 prosentilla ryhmään kuuluvista henkilöistä. (Taulukko 8). Puuttuva koulutustieto oli 40 prosentilla henkilöistä (Taulukko 12, liitteenä) eli lopulta 86 prosentilla klusteriin kuuluvista henkilöistä oli korkeintaan toisen asteen koulutus. 37 prosenttia klusteriin kuuluvista henkilöistä asui yksin (Taulukko 12, liitteenä).

Työssäolokuukausia oli viisi ja työttömyyskuukausia alle puoli kuukautta. Toimeentulotukea oli saatu keskimäärin yksi kuukausi vuosina 2010-2014. Perhe-etuudet olivat harvinaisia. (Taulukko 9).

Henkilöillä oli ollut keskimäärin yli 1600 päivää työkyvyttömyyseläkettä vuosina 2010-2014, eli he olivat olleet työkyvyttömyyseläkkeellä miltei koko ajanjakson. Määräaikaiset työkyvyttömyysetuudet olivat harvinaisia. Hoitoilmoituspäiviä eli erikoissairaanhoidon käyttöä oli 31 kappaletta. Sairauspäivärahopäiviä oli kerääntynyt keskimäärin noin 19 päivää. (Taulukko 10).

<b>KLUSTERIT MUUTTUJAT</b>	<b>1 (N=39 147)</b>	<b>2 (N=91 77)</b>	<b>3 (N=64 00)</b>	<b>4 (N=422 80)</b>	<b>5 (N=597 90)</b>	<b>6 (N=406 36)</b>	<b>7 (N=750 06)</b>	<b>8 (N=112 468)</b>	<b>9 (N=240 818)</b>
Naiset, %	48,6	57,2	72,4	48,4	57,7	7,0	19,7	88,1	74,5
Kaupunkilaiset, %	60,5	70,5	63,2	61,9	72,6	20,0	27,3	72,0	79,6
Ulkomaalaiset, %	0,30	1,50	0,10	0,40	1,80	1,20	0,50	2,80	1,90
0-4 vuotta Suomessa asuneet, %	0,0	0,2	0,0	0,0	0,1	0,1	0,0	0,1	0,1
Muun kuin suomen tai ruotsinkieliset, %	3,9	5,4	5,7	4,7	5,7	6,6	7,5	7,1	7,3
55 -64 -vuotiaat, %	60,6	8,5	69,2	19,7	38,9	50,2	5,5	79,3	0,0
Toisen asteen koulutus, %	46,0	57,4	46,5	37,6	67,2	14,1	62,7	18,4	20,3
Alimman korkea-asteen koulutus, %	9,0	8,4	17,5	9,2	9,6	6,9	9,6	12,4	16,8
AMK- tai kandidaatin tutkinto, %	2,7	8,0	4,4	3,0	5,6	2,5	4,8	4,3	7,2
Ylempi AMK- tai maisterin tutkinto, %	2,1	4,4	3,9	2,4	3,3	1,1	2,6	2,2	7,1
Eläkeläiset, %	96,0	54,4	58,5	96,4	16,4	77,3	96,9	63,9	91,8
Ei alle 18-vuotiaita lapsia, %	99,0	72,2	99,5	99,2	80,8	93,6	98,3	98,2	99,2
Opiskelijat, %	0,2	13,5	0,1	0,2	5,4	0,5	0,1	1,3	1,8
Ase- tai siviilipalveluksessa, %	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Asuntokunnassa 3 tai enemmän, %	8,4	31,9	7,9	8,3	27,8	14,1	11,0	5,2	8,0
Ikä, vuonna 2014 (Keskihajonta)	63,6 (3,2)	41,1 (9,9)	63,6 (2,0)	65,5 (3,9)	48,6 (11,1)	58,2 (10,2)	65,2 (7,4)	56,8 (9,5)	64,0 (10,9)

Taulukko 8 Työkyvyttömyyteen liittyviä etuuksia saanut väestö N=625 722. Yhdeksän ryhmän ratkaisu: ikä sekä taustamuuttujia, prosenttiosuus klusterissa. Sarakkeet edustavat ryhmiä, rivit kuvaavat ryhmien piirteitä eri taustatekijöittäin.

<b>KLUSTERIT MUUTTUJA T</b>	<b>1 (N=391 47)</b>	<b>2 (N=91 77)</b>	<b>3 (N=64 00)</b>	<b>4 (N=422 80)</b>	<b>5 (N=597 90)</b>	<b>6 (N=406 36)</b>	<b>7 (N=750 06)</b>	<b>8 (N=112 468)</b>	<b>9 (N=240 818)</b>
<b>Työttömyys, kuukaudet</b>	0,3 (2,7)	5,1 (8,1)	3,7 (11,6)	1,0 (4,4)	10,3 (13,9)	4,9 (10,7)	1,5 (5,6)	6,1 (13,7)	1,6 (6,2)
<b>Työssäolo, kuukaudet</b>	4,5 (12,0)	17,2 (19,8)	38,1 (21,6)	5,5 (12,5)	35,1 (22,4)	20,0 (22,2)	7,7 (14,1)	23,8 (25,0)	9,3 (15,2)
<b>Toimeentul otuki, kuukaudet</b>	1,0 (4,9)	7,1 (12,0)	0,4 (3,4)	1,1 (5,4)	6,1 (13,4)	2,3 (8,2)	0,4 (3,4)	2,3 (8,5)	0,9 (5,3)
<b>Perhe- etuudet, päivät</b>	0,3 (14,9)	25,1 (122,8)	0,0 (1,1)	0,2 (11,3)	4,4 (31,5)	0,6 (9,6)	0,1 (2,9)	0,3 (7,4)	0,1 (2,9)

Taulukko 9. Työkyvyttömyyteen liittyviä etuusia saanut väestö, N=625 722. Yhdeksän ryhmän ratkaisu: työttömyys, työssäolo ja ei-työkyvyttömyyteen liittyvät etuudet. Sarakkeet edustavat ryhmiä, rivit kuvaavat ryhmien piirteitä eri taustatekijöittäin.



KLUSTERIT MUUTTUJAT	1 (N=39 147)	2 (N=91 77)	3 (N=64 00)	4 (N=42 280)	5 (N=59 790)	6 (N=40 636)	7 (N=75 006)	8 (N=112 468)	9 (N=240 818)
Täysi työkyvyttömyysel äke, päivät	1629,0 (202,3)	0,1 (7,1)	0,3 (17,1)	830,1 (234,7)	4,6 (38,4)	14,1 (64,5)	23,0 (82,3)	2,3 (27,3)	10,0 (53,0)
Kuntoutustuki, päivät	0,2 (9,9)	1151,7 (413,2)	0,7 (14,3)	0,2 (6,7)	39,9 (112,4)	13,0 (67,5)	1,9 (25,5)	1,8 (25,1)	0,1 (3,7)
Osatyökyvyttömy yseläke, päivät	0,3 (18,7)	0,0 (0,0)	1250,9 (383,1)	0,2 (9,7)	1,1 (22,3)	1,2 (23,5)	2,1 (30,8)	1,8 (28,9)	2,4 (32,3)
Osakuntoutustuk i, päivät	0,0 (2,9)	1,1 (31,7)	0,1 (3,6)	0,1 (5,8)	13,4 (107,2)	2,3 (41,3)	0,5 (20,8)	0,8 (30,3)	0,0 (1,5)
Kuntoutusraha, päivät	0,2 (5,5)	21,8 (107,5)	2,5 (20,7)	0,3 (8,5)	61,6 (209,5)	4,4 (40,5)	1,1 (25,8)	0,9 (16,0)	0,0 (1,5)
Osakuntoutusav ustus, päivät	0,0 (0,0)	0,0 (0,3)	0,0 (0,0)	0,0 (0,0)	0,0 (0,0)	0,0 (0,0)	0,0 (0,0)	0,0 (0,2)	0,0 (0,0)
Osakuntoutusrah a, päivät	0,0 (0,0)	0,2 (9,3)	0,0 (0,0)	0,0 (0,0)	0,2 (11,8)	0,0 (0,0)	0,0 (0,0)	0,0 (0,0)	0,0 (0,0)
Kuntoutusavustu s, päivät	0,0 (0,0)	0,2 (8,5)	0,0 (0,0)	0,0 (0,0)	0,2 (9,0)	0,0 (1,5)	0,0 (0,3)	0,0 (3,5)	0,0 (0,0)
Kuntoutuskorotu s, päivät	0,2 (6,3)	47,1 (174,1)	0,8 (9,8)	0,1 (5,3)	4,3 (32,1)	1,3 (18,5)	0,2 (7,9)	0,5 (13,9)	0,0 (1,7)
Kelan kuntoutusraha, päivät	0,1 (2,5)	4,8 (38,0)	1,1 (6,2)	0,2 (4,6)	7,8 (52,2)	1,4 (19,2)	0,2 (7,5)	0,9 (16,6)	0,1 (3,0)
Hoitoilmoitukset, päivät	31,0 (67,1)	111,0 (134,0)	14,7 (26,3)	29,9 (62,0)	48,2 (81,9)	30,3 (63,8)	19,0 (50,4)	31,3 (80,7)	11,2 (27,5)
Sairauspäiväraha, päivät	19,4 (62,5)	144,9 (131,2)	24,0 (60,9)	36,5 (87,2)	247,8 (95,1)	117,8 (127,1)	18,8 (61,8)	14,4 (48,5)	2,1 (17,6)
Osasairauspäivär aha, päivät	0,1 (2,4)	1,5 (9,7)	1,5 (9,6)	0,2 (2,9)	4,5 (17,7)	1,2 (9,2)	0,2 (3,7)	0,6 (6,3)	0,0 (1,3)

Taulukko 10. Työkyvyttömyyteen liittyviä etuuksia saanut väestö, N=625 722. Yhdeksän ryhmän ratkaisu: työkyvyttömyyteen liittyvät etuudet. Sarakkeet edustavat ryhmiä, rivit kuvaavat ryhmien piirteitä eri taustatekijöittäin.

Toinen ryhmä, "kuntoutustuella olevat osatyökykyiset", koostui yhteensä vain 9177 henkilöstä, joista naisia oli 57 prosenttia. Keskimääräinen ikä oli 41 vuotta. Henkilöt olivat nuorempia kuin henkilöt työkyvyttömyysetuusia saaneiden väestössä keskimäärin. Henkilöistä 54 prosenttia oli eläkeläisiä ja 14 prosenttia opiskelijoita. Lisäksi 71 prosenttia henkilöistä asui kaupungissa ja kolmasosalla oli alle 18-vuotiaita lapsia. Toisen asteen koulutus oli 57 prosentilla ja vähintään alimman korkea-asteen koulutus hieman yli viidesosalla. (Taulukko 8). Viidesosalla puuttui koulutustieto kokonaan (Taulukko 12, liitteenä), joten henkilöt klusterista noin 80 prosenttia omasi korkeintaan toisen asteen koulutuksen. Henkilöistä 39 prosenttia asui yksin (Taulukko 12, liitteenä).

Työssäolokuukausia oli keskimäärin noin 17, työttömyyskuukausia viisi ja toimeentulotukikuukausia keskimäärin seitsemän. Lisäksi heillä oli muihin työkyvyttömyyteen liittyviä etuuksia saaneiden ryhmiin verrattuna eniten perhe-etuuspäiviä, keskimäärin noin 25 päivää. (Taulukko 9).

Toisen ryhmän henkilöt olivat saaneet keskimäärin 1152 päivää kuntoutustukea. Kuntoutuskorotusta oli saatu 47 päivää. Sairauspäivärahopäiviä oli kertynyt yhteensä keskimäärin 145 vuosina 2010-2014. Heillä oli muihin ryhmiin verrattuna paljon hoitoilmoituspäiviä, noin 111 päivää, eli heillä oli erikoissairaanhoidossa hoidettavia terveysongelmia. Osasairauspäivärahan käyttö oli yleisintä kaikista työkyvyttömyysetuuksia saaneiden ryhmistä, keskimäärin puolitoista päivää (Taulukko 10).

Kolmas ryhmä, "osatyökyvyttömyyseläkeläiset", koostui yhteensä 6400 henkilöstä, joista naisia oli 72 prosenttia. Henkilöiden keskimääräinen ikä oli 64 vuotta. Henkilöistä 59 prosenttia oli eläkeläisiä ja 63 prosenttia asui kaupungissa. Vain puolella prosentilla oli alle 18-vuotiaita lapsia. Lähes 50 prosentilla oli toisen asteen koulutus. Vähintään alimman korkea-asteen koulutus oli 26 prosentilla henkilöistä. (Taulukko 8). Puuttuvia koulutustietoja oli 27 prosentilla henkilöistä. Yksinasuvia oli hieman yli kolmasosa ryhmään kuuluvista henkilöistä. (Taulukko 12, liitteenä.)

Kolmannessa ryhmässä oli muihin ryhmiin verrattuna eniten työkuukausia, keskimäärin noin 38 kuukautta. Työttömyyskuukausia heillä oli neljä. Toimeentulotukikuukaudet ja perhe-etuudet olivat harvinaisia. (Taulukko 9).

Henkilöillä oli kertynyt keskimäärin 1251 päivää osatyökyvyttömyyseläkettä vuosina 2010-2014. Sairauspäivärahopäiviä oli kertynyt noin 24, osasairauspäivärahopäiviä noin puolitoista ja hoitoilmoituspäiviä noin 15. (Taulukko 10).

Ensimmäisen osatyökykyisten ryhmän tapaan, neljättä ryhmää voi kuvailla nimellä "työkyvyttömät eläkeläiset". Kaikkiaan 42 280 henkilöstä 48 prosenttia oli naisia. Henkilöiden keski-ikä oli 65 vuotta. Henkilöistä 96 prosenttia oli eläkeläisiä. 62 prosenttia henkilöistä asui kaupungissa. Hieman alle prosentilla ryhmään kuuluvista henkilöistä oli alle 18-vuotiaita lapsia. Toisen asteen koulutus oli 38 prosentilla ryhmään kuuluvista henkilöistä. Vähintään alimman korkea-asteen koulutus oli 15 prosentilla. (Taulukko 8). Puuttuva koulutustieto oli lähes puolella henkilöistä. Yli kolmasosa henkilöistä asui yksin. (Taulukko 12, liitteenä).

Työssäolokuukausia oli noin kuusi ja työttömyyskuukausia yksi. Toimeentulotukikuukausia ja perhe-etuuksia oli vain vähän. (Taulukko 9).

Henkilöillä oli kertynyt yli 800 päivää täyttä työkyvyttömyyseläkettä vuosina 2010-2014. Työkyvyttömyyseläkepäivien lisäksi sairauspäivärahopäivien keskiarvo oli 37 ja

hoitoilmoituspäivien 30. Määräaikaiset työkyvyttömyyden etuudet olivat harvinaisia, kun näiden etuuksien keskiarvot olivat lähellä nollaa. (Taulukko 10.)

Viides ryhmä, "osatyökykyiset keski-ikäiset", koostui 59 790 henkilöstä, joista naisia oli 58 prosenttia. Keskimääräinen ikä oli 49 vuotta. Henkilöistä 16 prosenttia oli eläkeläisiä ja viisi prosenttia opiskelijoita. 73 prosenttia henkilöistä asui kaupungissa. Noin viidesosalla ryhmään kuuluvista henkilöistä oli alle 18-vuotiaita lapsia. Viidennessä ryhmässä toisen asteen koulutuksen saaneita oli yli puolet, noin 67 prosenttia, ja korkeammin koulutettuja noin 19 prosenttia. (Taulukko 8). Puuttuva koulutustieto oli 13 prosentilla. 38 prosenttia klusteriin kuuluvista oli yksin asuvia. (Taulukko 12, liitteenä).

Työssäolokuukausia oli keskimäärin noin 35 kuukautta. Työttömyyskuukausia oli eniten muihin osatyökykyisten klustereihin verrattuna, keskimäärin noin 10 kuukautta vuosina 2010-2014. Vastaavasti toimeentulotuen tarve oli suurin, kuusi kuukautta. Perhe-etuuksia oli kertynyt keskimäärin neljä päivää. (Taulukko 9).

Henkilöillä oli keskimäärin noin viisi työkyvyttömyyseläkepäivää ja noin 40 kuntoutustukipäivää. Heillä oli kaikista klustereista eniten määräaikaisia kuntoutuspäiviä, noin 13, sekä kuntoutusrahopäiviä noin 61. Henkilöt olivat kerryttäneet keskimäärin 248 päivää sairauspäivärahaa ja noin neljä päivää osasairauspäivärahaa vuosina 2010-2014. Heillä oli myös muihin ryhmiin verrattuna melko paljon hoitoilmoituspäiviä, noin 48 päivää, eli heillä oli erikoissairaanhoidossa hoidettavia terveysongelmia. (Taulukko 10).

Kuudes ryhmä, "osatyökykyiset keski-ikäiset miehet", koostui 40 636 henkilöstä, joista naisia oli vain 7 prosenttia. Klusteriin kuuluvien keskimääräinen ikä oli 58 vuotta. Henkilöistä 77 prosenttia oli eläkkeellä. Vain 20 prosenttia henkilöistä asui kaupungissa, ja alle 7 prosentilla heistä oli alle 18-vuotiaita lapsia. Toisen asteen koulutus oli 14 prosentilta henkilöistä, korkeampi koulutus hieman alle kymmenesosalta. (Taulukko 8). Puuttuva koulutustieto oli 76 prosentilla henkilöistä. Hieman yli viidesosa asui yksin. Miltei kaksi kolmesta asui taajamassa eli kaupunkilaisten osuus oli pieni. (Taulukko 12, liitteenä).

Työssäolokuukausia oli keskimäärin noin 20 ja työttömyyskuukausia viisi. Toimeentulotukea oli saatu noin kahtena kuukautena vuosina 2010-2014. Perhe-etuudet olivat harvinaisia, keskimäärin puoli päivää. (Taulukko 9).

Kuudennessa klusterissa olevilla henkilöillä oli kertynyt täyttä työkyvyttömyyseläkettä keskimäärin noin 14 päivää ja kuntoutustukea noin 13 päivää.

Osatyökyvyttömyyseläkettä oli saatu keskimäärin yksi päivä, määräaikaista kuntoutustukea kaksi päivää, kuntoutusrahaa noin neljä päivää, kuntoutuskorotusta yksi päivä ja Kelan kuntoutusrahaa yksi päivä. Sairauspäivärahopäiviä oli kertynyt noin 118 päivää ja hoitoilmoituspäiviä noin 30 vuosina 2010-2014. (Taulukko 10).

Seitsemäs ryhmä, "vanhuuseläkeläiset", koostui 75 006 henkilöstä, joista 20 prosenttia oli naisia. Henkilöiden keski-ikä oli 65 vuotta eli he olivat jo ylittäneet virallisen eläkeiän. Henkilöistä 97 prosenttia oli eläkeläisiä. 27 prosenttia asui kaupungissa. Hieman alle kahdella prosentilla oli alle 18-vuotiaita lapsia. 63 prosentilla oli toisen asteen koulutus. Vähintään alimman korkea-asteen koulutus oli noin 17 prosentilla klusteriin kuuluvista henkilöistä. (Taulukko 8). Puuttuva koulutustieto oli noin viidenneksellä. 13 prosenttia asui yksin. (Taulukko 12, liitteenä).

Työssäolokuukausia oli noin kahdeksan ja työttömyyskuukausia kaksi. Toimeentulotukikuukausia oli kertynyt keskimäärin noin puoli. Perhe-etuudet olivat harvinaisia. (Taulukko 9).

Seitsemänteen ryhmään kuuluvilla henkilöillä oli keskimäärin 23 päivää täyttä työkyvyttömyyseläkettä vuosina 2010-2014. Määräaikaiset tai osa-aikaiseen työkyvyttömyyteen liittyvät etuudet olivat vähemmällä käytöllä, kun kuntoutustukea oli saatu noin kaksi päivää, määräaikaista kuntoutustukea noin puoli päivää. Lisäksi kuntoutusrahaa oli saatu noin yksi päivä. Sairauspäivärahopäiviä ja hoitoilmoituspäiviä oli kertynyt noin 19 kutakin. (Taulukko 10).

Kahdeksas ryhmä, "osatyökykyiset keski-ikäiset naiset", koostui 112 468 henkilöstä, joista 88 prosenttia oli naisia. Henkilöiden keski-ikä oli 57 vuotta. Henkilöistä 64 prosenttia oli eläkeläisiä. 72 prosenttia henkilöistä asui kaupungissa. Muihin työkyvyttömyyteen liittyviä etuuksia saaneiden ryhmiin verrattuna ryhmässä oli eniten ulkomaalaisia, noin 3 prosenttia. Hieman alle kahdella prosentilla henkilöistä oli alle 18-vuotiaita lapsia. Alle viidesosalla oli toisen asteen koulutus. 12 prosentilla oli alemman korkea-asteen koulutus. Sitä korkeampi koulutus oli noin seitsemän prosentilla henkilöistä. (Taulukko 8). Koulutusastetieto puuttui 62 prosentilla henkilöistä. Hieman yli viidesosa oli yksin asuvia. (Taulukko 12, liitteenä).

Työssäolokuukausia oli kertynyt noin 24, työttömyyskuukausia kuusi ja toimeentulotukikuukausia kaksi. Perhe-etuudet olivat harvinaisia. (Taulukko 9).

Henkilöillä oli kertynyt maltillisesti työkyvyttömyyteen liittyviä etuuksia: noin kaksi päivää täyttä työkyvyttömyyseläkettä, osatyökyvyttömyyseläkettä ja kuntoutustukea sekä noin yksi päivä määräaikaista kuntoutustukea, kuntoutusrahaa tai Kelan kuntoutusrahaa. Sairauspäivärahopäiviä oli kertynyt noin 14 ja hoitoilmoituspäiviä noin 31 vuosina 2010-2014. (Taulukko 10).

Seitsemännen ryhmän tapaan myös yhdeksäs ryhmä koostui eläkeiän ylittäneistä henkilöistä. 240 818 henkilöstä noin 75 prosenttia oli naisia. Henkilöiden keski-ikä oli 64 vuotta. Henkilöistä 92 prosenttia oli eläkeläisiä. Hieman alle prosentilla henkilöistä oli alle 18-vuotiaita lapsia. Toisen asteen koulutus oli joka viidennellä, alimman korkea-asteen koulutus 16 prosentilla, ammattikorkeakoulu- tai yliopistotutkinto noin 14 prosentilla. (Taulukko 8.). Tutkijakoulutettuja oli yksi prosentti, koulutusastetieto puuttui lähes 48 prosentilta. Henkilöistä 29 prosentti asui yksin. (Taulukko 12, liitteenä).

Työssäolokuukausia oli noin yhdeksän ja työttömyyskuukausia kaksi. Toimeentulotukea oli saatu keskimäärin noin yksi kuukausi. Perhe-etuudet olivat harvinaisia. (Taulukko 9).

Henkilöillä oli kertynyt noin 10 työkyvyttömyyseläkepäivää ja noin kaksi osatyökyvyttömyyseläkepäivää. Muutoin määrä- tai osa-aikaiset työkyvyttömyyteen liittyvät etuudet olivat harvinaisia. Sairauspäivärahopäiviä oli kertynyt vain kaksi ja hoitoilmoituspäiviä 11. (Taulukko 10).

## 5 Pohdinta ja rajoitukset

### 5.1 Pohdinta

K-means-ryhmittelymenetelmän kategoristen muuttujien ryhmittelyyn soveltuva variaatiota hyödyntäen löydettiin kymmenen työhön osallistumiseltaan ja työkyvyltään toisistaan eroavaa työikäisten ryhmää. Tarkasteltavana oli 90 prosentin otos työikäisestä väestöstä. Ryhmien ominaisuuksia ja eroavaisuuksia kartoitettiin tarkastellen eri muuttujien ryhmäkohtaisia keskiarvoja tai osuuksia. Näin 90 % väestöotoksen ryhmittelyn tuloksena löydettiin yhteensä viisi ryhmää, joissa oli hyödynnetty muita enemmän työkyvyttömyyteen liittyviä etuuksia. Näihin ryhmiin kuuluvat yli 600 000 henkilöä poimittiin jatkotarkasteluun, jossa toistettiin sama analyysi uudestaan pienemmälle väestölle. Tässä pohdinnassa käydään läpi tuloksia ja niiden hyödynnettävyyttä. Lisäksi huomiota kiinnitetään aineistoon ja menetelmiin liittyviin rajoituksiin sekä mahdollisiin tulevaisuudessa toteutettaviin tarkasteluihin.

Miten tulosta tulisi tulkita? 90 prosentin väestöotoksen ryhmittelyratkaisusta ei löytynyt täysin työelämän ulkopuolella olevien ryhmää, joka ei olisi eläkkeellä. Toisin sanoen, esimerkiksi pitkäaikaistyöttömät eivät erottuneet ryhmittelyratkaisussa omana ryhmänään. Yksi löytynyt työelämän ulkopuolella olevien ryhmä oli täydellä työkyvyttömyyseläkkeellä olleet henkilöt. Tämä voi johtua esimerkiksi käytettyjen etuustietojen painottumisesta työkyvyttömyyteen liittyviin etuuksiin. Eli työkyvyttömyyteen liittyvää etuustietoa hyödynnettiin enemmän kuin työttömyyteen liittyvää etuustietoa. Tarkastelua voi rajata sekä erilaisiin väestönosiin tai toisaalta tarkastelussa voisi rajoittaa käytetyn tiedon määrää. Mikäli tarkastelun rajaisi esimerkiksi työttömyyskuukausia omanneisiin henkilöihin, voitaisiin löytää toisistaan eroavia työttömien ryhmiä. Tulosten vertailu Faces of joblessness -tutkimukseen (OECD, 2020) voisi tarjota lisätietoa heikosti työllistyvien väestön haasteista ja rekisteritiedon hyödyntämisen mahdollisuuksista tämän väestönosan tarkastelussa. Tässä tutkimuksessa otettiin jatkotarkastelun kohteeksi työkyvyttömyyden etuuksia saaneiden väestö.

Työkyvyttömyyteen liittyviä etuuksia saaneiden väestöön rajattua otosta käyttäen saatiin tarkentuva kuvaus työkyvyttömyyden etuuksia saaneista henkilöistä. Tarkempi kuva perustui yhdeksään toisistaan eroavaan ryhmään. Ryhmiä erotteli se, kuinka pitkäkestoista tai kuinka laajaa työkyvyttömyyteen liittyvien etuuksien käyttö oli ollut. Tämä näkyi esimerkiksi pitkäkestoisen työkyvyttömyyseläkkeen ja määräaikaisten työkyvyttömyyteen liittyvien etuuksien käytön eroina ryhmissä. Ryhmät erosivat

toisistaan myös työssäolo-, työttömyys- ja toimeentulotukikuukausien mukaisesti. Lisäksi ne erosivat toisistaan esimerkiksi iän, sukupuolen ja asutokunnan koon mukaan.

Tarkentuva tulos rohkaisee kokeilemaan ryhmittelymenetelmien hyödyntämistä myös muiden, kuten esimerkiksi työttömyyskuukausia kerryttäneiden henkilöiden tarkasteluun. Ryhmittelyanalyysi tarjoaa lisää tietoa erilaisista väestönosista jakaen väestönosat pienempiin ryhmiin, jotka eroavat toisistaan tavalla tai toisella. Tässä toteutettu ryhmittely perustui taustatietojen lisäksi etuustietoihin ja tietoihin erikoissairaanhoidon käytöstä. Etuustiedot ovat hallinnollisiin tarpeisiin kerättyä tietoa, mitä voidaan pitää erittäin luotettavana. On kuitenkin toisen tutkimuksen aihe, kuinka paljon etuuksien käytöstä voidaan päätellä esimerkiksi syy-seuraussuhteita vaikkapa työkyvyn ja työssäolon välisestä yhteydestä. Tämän tutkimuksen tarkoituksiin kytkeytyen voidaan todeta, että ne kertovat paljon työhön osallistumisesta, työtulojen sijaan saatavasta toimeentulosta ja täten ainakin mahdollisista syistä sille, miksi töihin ei osallistuta.

## 5.2 Tulosten hyödynnettävyys

Tietoa työhön osallistumisesta ja työtulojen sijaan saatavasta toimeentulosta voidaan hyödyntää päätöksenteon tukena suunniteltaessa työn tarjontaan liittyviä uudistuksia. Ensinnäkin ajantasaisilla tiedoilla toteutetulla ryhmittelyllä voidaan tunnistaa toisistaan eroavia ryhmiä, joita mahdolliset uudistukset koskevat eri tavoin. Ryhmien tunnistamisessa voidaan käyttää rajattua aineistoa koko väestöstä. Esimerkiksi henkilöt, joilla oli työttömyyskuukausia tai henkilöt, jotka saivat erikoissairaanhoidoa jollain aikavälillä. Käytännössä väestötason luotettavien etuustietojen hyödyntäminen mahdollistaa lukemattomien eri alaryhmien tunnistamisen erilaisia päättelyn tai päätöksenteon tarpeita varten.

Ryhmittely tarjoaa päätöksentekijälle mahdollisuuden arvioida

- kuinka suuri on kohdejoukko, jota suunniteltu uudistus koskee
- millaisia eri alaryhmiä kohdejoukkoon kuuluu
- miten alaryhmät eroavat toisistaan
- millaisia uudistuksen oletetut vaikutukset voivat olla eri alaryhmille
- millaisia uudistuksen oletetut vaikutukset voivat olla ryhmille, jotka eivät ole uudistuksen ensisijaisena kohteena

Edellä mainituista kohdista jokainen tukee mahdollisuutta parempaan etukäteen suoritettuun vaikutusten arviointiin. Ryhmittelyn tuloksena syntyviä tietoja voidaan

käyttää vaikutusten potentiaalin arviointiin. Tässä arvioinnissa voidaan ottaa huomioon kohdejoukko eri osajoukkoineen ja toisaalta myös ryhmiä, jotka eivät ole uudistuksen alkuperäisenä kohteena. Uudistuksen aiempaa tarkempi määrittely etukäteen kytkee uudistusten toteutumisen eri näkökulmiin. Yksi relevantti näkökulma on tarkentuva ennuste uudistuksen vaikutuksista työvoimapotentialiin eli työssäolokuukausien lisäämiseen. Tästä on tarjottu esimerkki työkyvyttömyysetuuksia hyödyntäneiden henkilöiden ryhmittelyratkaisua koskevassa pohdinnan osassa.

Ryhmittelyä voidaan hyödyntää myös vaikutusten arvioinnissa toteutuneen uudistuksen jälkeen. Tällöin ryhmittelyn kautta voidaan tunnistaa mahdollisesti toisistaan eroavia ryhmiä, joille uudistuksesta koitui mahdollisesti toisistaan eroavia vaikutuksia. Tämä edesauttaa politiikka-arvioinnissa sekä seuraavien uudistusten vaikutusten ennakoarvioinnissa ja suunnittelussa.

### 5.3 Työkyvyttömyyteen liittyviä etuuksia saaneen väestön ryhmittely

Koko väestöön verrattuna työkyvyttömyyteen liittyviä etuuksia saaneiden keskimääräinen koulutusaste oli matalampi. Heidän joukossaan oli pienempiä osuuksia ulkomailta Suomeen muuttaneita. Henkilöt olivat hieman iäkkäämpiä, ja siksikin heillä oli harvemmin alle 18-vuotiaita lapsia kotona. Useassa eri ryhmittelyratkaisussa heistä suuri osa oli eläkkeellä. Toimeentulotukea käytettiin jonkun verran, ja sen käyttö näytti olevan yleisempää nuorempien henkilöiden joukossa. Ryhmissä oli jonkin verran päätoimisia opiskelijoita.

Työkyvyttömyyden etuuksia hyödyntäneiden väestössä oli useita ryhmiä, joissa työssäolokuukaudet olivat yleisiä. Vähiten työssäolokuukausia löytyi ryhmistä yksi, neljä, seitsemän ja yhdeksän. Näissä ryhmissä työssäolokuukausien määrä oli viiden ja kymmenen kuukauden välillä vuosina 2010-2014.

Ensimmäinen ja neljäs klusteri muodostuivat ikääntyneistä henkilöistä, joista monet olivat olleet vuosina 2010-2014 täydellä työkyvyttömyyseläkkeellä. Ero ryhmien välillä oli se, että neljännen klusterin henkilöillä oli keskimäärin 830 työkyvyttömyyseläkepäivää vuosina 2010-2014, kun ensimmäisessä ryhmässä työkyvyttömyyseläkepäiviä oli yli 1600. Keskimääräinen koulutusaste oli hyvin matala molemmissa ryhmissä, kun lähes 90 prosenttia henkilöitä omasi korkeintaan toisen asteen koulutuksen. Yksi selitys pienemmälle työkyvyttömyyseläkkeen määrälle oli se,



että henkilöt neljännessä ryhmässä olivat muutaman vuoden ryhmän yksi henkilöitä vanhempia. Tämä tarkoittaa, että vuosina 2010-2014 he ovat ehtineet olla pidempään vanhuuseläkkeellä ja täten vähemmän työkyvyttömyyseläkkeellä.

Seitsemännen ryhmän osalta merkittävin ero työkyvyttömiä eläkeläisten ryhmiin yksi ja neljä oli se, että harvempi heistä asui kaupungissa. Ryhmä muodostui iän puolesta vanhuuseläkkeelle oikeutetuista henkilöistä, joista osalla oli kuitenkin ilmennyt työkyvyn ongelmia viimeisinä vuosina ennen vanhuuseläkettä. Lisäksi henkilöillä oli keskimäärin alle 25 työkyvyttömyyseläkepäivää, kun esimerkiksi ensimmäisessä työkyvyttömyyden etuuksia saaneiden väestön ryhmässä työkyvyttömyyseläkepäiviä oli yli 1600. Tilapäisten tai osa-aikaisten työkyvyttömyyden etuuksien saanti olivat yleisempiä kuin ryhmässä yksi ja neljä. Ryhmään kuuluvilla oli näitä kahta muuta ryhmää useammin toisen asteen koulutus.

Yhdeksäs ryhmä oli suurin tässä tarkastelussa muodostunut ryhmä, joka oli kooltaan jopa neljännesmiljoona henkilöä. Miltei kolme neljästä henkilöstä oli naisia. Miltei 90 prosenttia oli vähintään 65-vuotiaita. Täyttä työkyvyttömyyseläkettä oli saatu keskimäärin kymmenen päivää, osatyökyvyttömyyseläkettä kaksi päivää. Muut työkyvyttömyyteen liittyvät etuudet olivat harvinaisia.

Klustereissa kaksi ja kuusi oli noin 17-20 kuukautta työssäoloa vuosina 2010-2014.

Toisessa klusterissa oli yli 9000 henkilöä. Ryhmään kuului keskimäärin muita muodostuneita ryhmiä nuorempia henkilöitä. Keskimääräinen ikä oli 41 vuotta. Päätoimisia opiskelijoita oli 14 prosenttia. Työssäolokuukausia oli 17. Henkilöillä oli muihin ryhmiin verrattuna enemmän kuntoutustuen käyttöä, kun taas täysi, pysyvä työkyvyttömyys oli harvinaista saatujen etuuksien valossa tarkasteltuna. Kuntoutustuen käytön lisäksi ryhmässä oli paljon hoitoilmoituspäiviä, 111 kappaletta. Myös sairauspäivärahopäivät olivat yleisiä, kun näitä oli keskimäärin 145 kappaletta. Henkilöiden työkyvyn ongelmat voivatkin olla lähtöisin terveyden ongelmista, joita on hoidettu erikoissairaanhoidossa.

Kuudes ryhmä koostui suurilta osin miehistä, jotka eivät asuneet kaupungeissa. Heitä oli kaiken kaikkiaan yli 40 000. Keskimääräinen ikä oli 58 vuotta. Työssäolokuukausia oli 20. Henkilöillä oli orastavaa työkyvyttömyyttä, joka ilmeni sekä täyteen työkyvyttömyyteen että osatyökyvyttömyyteen liittyvien etuuksien maltillisena saamisena. Heillä oli keskimäärin paljon sairauspäivärahopäiviä, 117 kappaletta. Muita etuuksia ei oltu saatu keskimääräistä enempää verrattuna muihin ryhmiin, joskin työttömyyttä oli ollut keskimäärin viisi kuukautta. Voi olla, että harvenevat

työmahdollisuudet ja työttömyys ovat osaltaan vaikuttaneet myös työkyvyn rappeutumiseen.

Eniten työssäolokuukausia oli ryhmissä kolme, viisi ja kahdeksan, noin 24 kuukaudesta yli 38 kuukauteen.

Viidennessä ryhmässä oli noin 60 000 henkilöä, jotka olivat keskimäärin 49-vuotiaita. He olivat verrattain matalasti koulutettuja. Opiskelijoita oli noin viisi prosenttia. Työssäolokuukausia oli toiseksi eniten kaikista ryhmistä, 35 kuukautta. Lisäksi viidennessä ryhmässä oli enemmän työttömyyteen liittyvien etuuksien saantia kuin muissa ryhmissä. Tämä näkyi myös korkeahkona toimeentulotukikuukausien määränä. Näitä oli keskimäärin kuusi. Viidennessä ryhmässä ei ollut saatu täyden työkyvyttömyyden etuuksia. Henkilöillä oli keskimäärin pitkään jatkunutta sairauspäivärahan käyttöä, 247 päivää, ja noin 48 hoitoilmoituspäivää, joten heillä saattoi olla orastavaa tarvetta työkyvyttömyyteen liittyville ratkaisuille ja etuuksille. Henkilöt olivat vielä työikäisiä, joilla työkyvyn ongelmat voivat olla ensisijaisesti terveyden ongelmia, joita on hoidettu erikoissairaanhoidossa.

Kolmas ryhmä koostui 6400 henkilöstä, eli se oli hyvin maltillisen kokoinen ryhmä verrattuna muihin. Henkilöt olivat keskimäärin 64-vuotiaita. Yli 70 prosenttia oli naisia. Työssäolokuukausien määrä oli korkein kaikista tarkastelussa muodostuneista ryhmistä, 38 kuukautta. Henkilöt hyödynsivät usein lievempään työkyvyttömyyteen liittyvää etuutta, osatyökyvyttömyyseläketä. Hoitoilmoituspäivien ja sairauspäivärahopäivien taso oli huomattavasti esimerkiksi toista ryhmää vähempää. Muiden etuuksien käyttö oli harvinaista, mistä voidaan päätellä, että osatyökyvyttömyyseläke palveli tarkoitustaan siten, että se hillitsi terveyden ja työkyvyn ongelmia. Ainakin etuuksien saannin kautta tarkasteltuna.

Kahdeksannessa ryhmässä oli kaiken kaikkiaan 112 000 henkilöä, joista miltei 90 prosenttia oli naisia. Heidän keskimääräinen ikänsä oli 57 vuotta. 64 prosenttia oli eläkkeellä. Työssäolokuukausia oli toiseksi eniten tarkasteltavista klustereista, miltei 24 kappaletta. Työttömyyttä oli keskimäärin kuusi kuukautta. Ryhmässä oli myös korkein ulkomaalaistaustaisten osuus kaikista työkyvyttömyyden etuuksia saaneiden ryhmistä, miltei kolme prosenttia. Henkilöillä oli maltillista työkyvyttömyyteen liittyvien etuuksien saamista sekä täyden että osa- tai määräaikaisen työkyvyttömyyden varalle. Hoitoilmoituspäiviä oli yli 30, sairauspäivärahopäiviä vain 14, mikä oli vähemmän kuin useissa muissa ryhmissä. Hoitoilmoituspäiviä oli 11 ja sairauspäivärahopäiviä vain kaksi.

## 5.4 Skenaariolaskelman hahmottelua

Oletetaan politiikkatoimenpide, jonka tavoitteena on työn tarjonnan (tai työssäolokuukausien) lisääminen. Useat työkyvyttömyyteen liittyviä etuuksia saaneiden väestön tarkastelussa muodostuneet ryhmät kuuluvat potentiaaliseen kohderyhmään.

Ensinnäkin vanhuuseläkeiän ohittaneiden henkilöiden ryhmät voivat lisätä työvoiman tarjontaa, mikäli heillä on työkykyä. Tässä tarkastelussa keskitytään kuitenkin niihin, jotka ovat vanhuuseläkeikää nuorempia, ja ovat saaneet osa- ja tai määräaikaisia työkyvyttömyyteen liittyviä etuuksia. Tällöin tarkastelun kohteena ovat luvun 4.4 tarkastelussa muodostetut ryhmät kaksi, viisi, kuusi ja kahdeksan, joihin kuului yli 220 000 henkilöä.

Ryhmiiin kaksi, viisi, kuusi ja kahdeksan kuuluvien henkilöiden keskimääräiset iät olivat 41, 49, 58 ja 57 vuotta. Työssäolokuukaudet (ja työttömyyskuukaudet) vuosina 2010-2014 olivat ryhmittäin 17 (5), 35 (10), 20 (5), 24 (6). Ryhmissä oli saatu työkykyyn liittyvien etuuksien osalta kuntoutustukea, pitkiä sairauspäiväraajaksoja ja vähän myös täyttä työkyvyttömyyseläkettä. Lisäksi henkilöillä oli pitkittyvää työttömyyttä. Mikäli työkyvyttömyys olisi väliaikaista, tai jäljellä olevaa työkykyä voisi hyödyntää työmarkkinoilla ja henkilöille löytyisi työmahdollisuuksia, voisi vertailukelpoisen, seuraavan viisivuotiskauden keskimääräiseksi työssäolokuukausien määräksi tavoitella 35 kuukauden tasoa, mikä vastaa ryhmän viisi tasoa vuosina 2010-2014. Tämä tarkoittaisi noin 30 000 henkilötyövuoden lisäystä talouteen vuosittain.

## 5.5 Aineistoon liittyviä rajoituksia

Käytössä ollut aineisto on verrattain vanhaa. Seurantamme rajoittuu vuoteen 2014, josta on kulunut tätä kirjoitettaessa kahdeksan vuotta. Kaikki aineiston tiedot perustuvat myös yli 7 vuoden takaisin tietoihin rekistereistä, joten tuloksia ei voida yleistää koskemaan nykytilannetta. Tulokset menetelmän käytettävyydestä puoltavat analyysin toistamista uudemmilla ja jopa kattavammilla aineistoilla.

Toiseksi aineiston suuri koko asetti rajoituksia menetelmien hyödyntämiselle. Esimerkiksi ryhmittelyratkaisujen validointia ei voinut suorittaa kuin pienellä otoksella käytössä olevasta aineistosta. Lisäksi koko väestön hyödyntäminen ryhmittelyssä aiheutti hitautta laskentaan. Toisaalta ryhmittelyratkaisu väestön kattavilla aineistoilla voisi tuottaa enemmänkin ryhmiä, mutta jo kymmenen ryhmän ratkaisu on laskuteknisesti raskas toteutettava.

Kolmanneksi aineistossa käytetyt sairauspoissaolotiedot koskevat vähintään 10 päivän poissaolojaksoja, jolloin lyhyemmät sairauspoissaolot jäävät huomioimatta. Lyhyempien sairauspoissaolojen lisääminen analyysiin ei kuitenkaan ole mahdollista, koska rekisteriaineistoja niistä ei ole olemassa. Lisäksi tässä tarkastelussa hyödynnetyt sairauspäivärahatiedot olivat niitä, jotka kertyvät käytössä olleisiin rekistereihin vain siinä tapauksessa, että etuus on maksettu työntekijälle. Usein käytännössä sairauspäivärahaetuus maksetaan työnantajalle, joka jatkaa palkanmaksua sairaudesta huolimatta.

Neljänneksi pääasiallinen toiminta on tieto, joka sisältää vuoden pisimmän toiminnan. Esimerkiksi tieto eläkeläisten osuudesta on otettu pääasiallisesta toiminnasta, ja voi tarkoittaa mitä vain myönnettyä eläkettä. Tämä vaikeuttaa ryhmätason tulosten tarkastelua, mutta ongelma on korjattavissa tarkemmalla eläkeperusteiden tarkastelulla.

## 5.6 Menetelmään liittyviä rajoituksia

Yksi K-Prototypes-ryhmittelymenetelmän rajoitus on, että menetelmän käyttö ei ole mielekäästä suuresti poikkeavien havaintojen kohdalla. Kuten muissakin kovissa ryhmittelymenetelmissä, jokainen havaintoyksikkö jaetaan yhteen ryhmään, vaikka ryhmä ei olisi havaintoyksikölle sopiva. Kovat ryhmittelymenetelmät ovat kuitenkin laskennallisesti tehokkaampia kuin niin sanotut sumeat ryhmittelymenetelmät, joissa havainnoille lasketaan jokaiseen ryhmään kuulumisen todennäköisyys. Tässä tapauksessa aineiston koko ja käytettävissä oleva laskentateho johti kovien ryhmittelymenetelmien käyttöön.

Kovissa ryhmittelymenetelmissä tulee valita ryhmien määrä etukäteen. Tässä tutkimuksessa haluttu ryhmien määrä validoitiin käyttäen yhden prosentin satunnaisotosta aineistosta. Luotettavuus voisi olla parempi, mikäli validointiin kyettäisiin käyttämään suurempaa satunnaisotosta. Tämän tutkimuksen yhteydessä ei kyetty suorittamaan ryhmien sisäisten etäisyyksien laskentaa eikä validointiin käytettävien indeksien, kuten esimerkiksi Calinski-Harabasz-indeksin laskentaa lähes koko väestön kattavalla aineistolla.

## 6 Seuraavista askeleista

Tässä tutkimuksessa tarkasteltiin työkyvyttömyyttä ja työhön osallistumista etuuksien käytön näkökulmasta. Tämä tutkimus jättää useita mahdollisuuksia seuraaviksi askeliksi. Ensinnäkin tässä toteutettua tarkastelua voitaisiin jatkaa. Toisaalta tutkimuksen teeman eli työkyvyttömyyden ja työhön osallistumisen teemaa voitaisiin tutkia hyödyntäen tässä käytettyjä menetelmiä.

Yksi oleellinen lisä tähän tarkasteluun olisi jatkaa tarkastelua kartoittaen ryhmittelyanalyysissa määräytyneeseen ryhmään kuulumista selittäviä tekijöitä. Tätä voitaisiin tutkia muun muassa regressiomenetelmiä hyödyntäen. LASSO-regression avulla voitaisiin päätellä, mitkä muuttujat vaikuttavat eniten havaintojen ryhmittäytymiseen tiettyyn ryhmään. Multinomiaalisella logistisella regressioanalyysillä voitaisiin vertailla muuttujien vaikutusta ryhmittelyyn eri ryhmien välillä.

Toinen tähän tarkasteluun liittyvä lisä voisi olla ryhmien välisen liikehdinnän tarkastelu. Ryhmään kuuluminen ei ole olettavasti pysyvä tila, vaan ajan myötä yksilön tilanne kehittyy. Tämän ryhmien välisen liikehdinnän tarkastelu voisi tuoda lisäarvoa politiikkatoimenpiteiden arviointiin. Tämä edellyttäisi seuranta-aineistolla toteutettua ryhmittelytarkastelujen jatkumoa ja yksilöiden poimimista ryhmiin eri vuosille rakennetuista ryhmittelyratkaisuista.

Tähän tutkimukseen käytetyt aineistot vaativat päivitystä, jotta luotettavampaa tietoa työikäisen väestön nykytilanteesta voitaisiin tuottaa. Nyt tiedot perustuvat kahdeksan vuoden takaisiin rekisteritietoihin, joiden ei voida olettaa kuvaavan nykyhetken kaltaista tilannetta edes väestörakenteen osalta. Jatkossa aineistot voidaan päivittää ajallisesti vastaamaan lähempänä nykyhetkeä olevaa ajankohtaa. Aineistoa voidaan päivittää myös sisällöllisesti eli sisällyttää aineistoon uusia rekisteritietoja, kuten esimerkiksi terveyden osalta Kelan lääkeostot tai työttömyyden osalta Työ- ja elinkeinoministeriön alaiset koulutukset ja toimenpiteet. Uusien rekisteritietojen perusteella voidaan myös jatkaa analyysijä, ja esimerkiksi tutkia työkyvyttömyyttä aiheuttavien sairauksien määriä työkyvyttömiä klustereissa.

Näkökulmaa määrittävät verrattain helposti tutkimuskäyttöön saatavilla olevat etuusrekisteritiedot. Jatkossa olisi perusteltua tarkastella koko työikäisen väestön työkykyä hyödyntäen laajemmin muutakin tietoa kuin etuustietoja. On todennäköistä, että työkyky ja työkyvyttömyys vaihtelee ajassa ja osa työkyvyn alenemaan kokevista ei hakeudu palveluiden saatikka etuuksien piiriin. Tätä työtä jatketaan osana Euroopan

unionin rahoittamaa ohjelmaa "Suomen kestävä kasvun ohjelma – Työkykyohjelman laajennos (RRP)". Ohjelman tutkimusosiossa muodostetaan arviointikehikko työhön osallistumisesta ja työkyvystä sekä työkyvyn tukeen liittyvistä palveluista ja etuuksista. Arviointikehikko operationalisoidaan osa-alueiden rekisteriaineistojen avulla huomattavasti tämän tutkimuksen tietosisältöjä laajemmin.

## 7 Lopuksi

Tämän lyhytraportin tavoitteena oli tunnistaa saatujen työkykyyn liittyvien etuuksien osalta – ja täten työkyvyltään – toisistaan eroavia ryhmiä. Tämä tehtiin aineistolähtöisesti hyödyntäen ryhmittelymenetelmiä lähes koko väestön kattavalla aineistolla. Tuloksena oli kymmenen toisistaan eroavaa ryhmää, joista viidessä oli saatu erinäisiä työkyvyttömyyteen liittyviä etuuksia. Ryhmittely suoritettiin uudelleen näille viidelle työkyvyttömyyteen liittyviä etuuksia saaneiden ryhmälle. Näin saatiin entistä tarkempi ryhmittelytulos työkyvyttömyyteen liittyviä etuuksia saaneiden henkilöistä.

Tämän tutkimuksen tulokset osoittavat, että tarkentavaa ryhmittelyä voitaisiin toteuttaa myös muiden työhön osallistumiseen liittyvien ryhmätasoisien piirteiden perusteella.

Lopulta koko työikäinen väestö voidaan ryhmitellä pienempiin osiin hyödyntäen erinäisiä tietoja, kuten tässä tutkimuksessa on hyödynnetty erityisesti työssäolo- ja etuustietoja. Näin saadaan jäsennellympää tietoa väestöstä ja väestön jakautumisesta erilaisiin ryhmiin. Tätä tietoa voidaan hyödyntää näitä ryhmiä koskevan päätöksenteon suunnittelun ja arvioinnin tukena.

## 8 Lähteet

Halkidi, M., Batistakis, Y., & Vazirgiannis, M. (2001). On clustering validation techniques. *Journal of intelligent information systems*, 17(2), 107-145.

Huang, Z. (1998). Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data mining and knowledge discovery*, 2(3), 283-304.

OECD (2020). Faces of Joblessness in Finland – A people-centred Perspective on Employment Barriers and Policies. Saatavana:  
[https://valtioneuvosto.fi/documents/1271139/2013549/Finland+FOJ\\_FINAL+updated+9+October+2020.pdf/1b0913b3-5222-3e97-dd61-de4e33333e11/Finland+FOJ\\_FINAL+updated+9+October+2020.pdf?t=1602238273788](https://valtioneuvosto.fi/documents/1271139/2013549/Finland+FOJ_FINAL+updated+9+October+2020.pdf/1b0913b3-5222-3e97-dd61-de4e33333e11/Finland+FOJ_FINAL+updated+9+October+2020.pdf?t=1602238273788)

Työvoimatutkimus (2022). Työvoimatutkimus 2022, helmikuu. Tilastokeskus. Viitattu 22.11.2022. Saatavana <http://www.stat.fi/til/tyti/index.html>

Valtioneuvosto (2021). Suomi kohti kestäväää kasvua. Valtioneuvosto. Viitattu 10.9.2021. Saatavana  
<https://valtioneuvosto.fi/documents/10184/1058456/Valtiovarainministeri+Saarikon+esitysmateriaali+9.9.2021.pdf/866cf7e9-aea2-1f6b-300b-e743d0c9f396/Valtiovarainministeri+Saarikon+esitysmateriaali+9.9.2021.pdf?t=1631195136935>



## 9 Liitteet

### 9.1 Liitetaulukot

KLUSTERIT MUUTTUJAT	1 (N=408 45)	2 (N=932 2)	3 (N=427 73)	4 (N=895 26)	5 (N=4432 56)	6 (N=1329 24)	7 (N=5098 41)	8 (N=7646 06)	9 (N=5615 90)	10 (N=5730 49)
Asuinpaikka:	18,7%	16,3%	19,2%	18,5%	17,4%	17,1%	18,0%	14,9%	12,8%	15,7%
taajama										
maaseutumaiset	20,8%	13,4%	20,1%	18,7%	16,9%	13,5%	14,7%	13,1%	11,1%	13,3%
Kansalaisuus:	99,6%	98,6%	99,6%	98,1%	98,2%	94,7%	96,9%	97,3%	97,8%	96,7%
suomi										
Maassaoloaika:	0,1%	0,4%	0,2%	0,3%	0,2%	0,6%	0,4%	0,4%	0,3%	0,5%
5-7 vuotta maassa										
8-10 vuotta maassa	0,1%	0,5%	0,2%	0,3%	0,2%	0,7%	0,5%	0,3%	0,3%	0,5%
yli 10 vuotta maassa	0,8%	2,0%	0,8%	1,7%	0,9%	1,6%	2,0%	1,0%	1,3%	1,9%
Kieli:	95,9%	94,0%	95,1%	93,1%	91,9%	87,2%	89,1%	91,1%	91,8%	88,8%
suomi										
ruotsi	0,1%	0,5%	0,2%	0,8%	0,9%	1,9%	1,5%	1,2%	1,1%	1,1%
Ikä luokiteltuna:	0,0%	1,5%	0,0%	1,6%	2,3%	2,9%	4,3%	2,6%	40,8%	0,2%
alle 24-vuotiaat										
24-34 -vuotiaat	0,3%	29,0%	0,3%	10,1%	2,0%	52,1%	15,5%	48,4%	8,9%	8,2%
35-44 -vuotiaat	0,4%	29,0%	0,7%	14,2%	2,7%	41,7%	33,3%	9,2%	14,1%	28,2%
45-54 -vuotiaat	0,9%	31,4%	0,8%	26,2%	5,8%	3,2%	45,3%	18,7%	18,8%	18,3%
yli 65 -vuotiaat	40,0%	0,1%	77,9%	2,7%	64,6%	0,0%	0,1%	2,0%	3,5%	1,0%
Lasten lukumäärä:	0,7%	14,1%	0,6%	8,2%	0,8%	18,2%	50,9%	4,0%	3,3%	9,5%
Yksi alle 18-										
vuotias lapsi										
2-3 alle 18-	0,2%	12,3%	0,3%	7,6%	0,2%	68,2%	37,5%	4,2%	1,8%	27,8%
vuotiasta lasta										
4 tai enemmän alle 18-vuotiaasta lasta	0,0%	1,3%	0,0%	0,8%	0,0%	12,5%	2,8%	0,4%	0,3%	2,4%
Asutokunnan koko:	37,0%	39,0%	34,3%	28,6%	24,5%	0,6%	0,1%	11,3%	70,1%	13,5%
Asutokunta 1										
Asutokunta 2	54,7%	29,0%	57,3%	46,6%	67,8%	4,6%	5,6%	75,4%	11,8%	22,2%
Koulutus:	0,2%	0,5%	0,2%	0,9%	0,2%	0,5%	1,1%	0,8%	0,4%	1,6%
erikoisammattitut kinto										
lisensiaatti tai tohtorin tutkinto	0,2%	0,2%	0,2%	0,3%	0,7%	0,7%	1,5%	0,7%	0,5%	2,1%
puuttuvat	41,1%	21,3%	45,7%	23,6%	49,0%	13,3%	8,2%	8,0%	14,5%	16,1%

Taulukko 11. 90 % väestöotos, N=3 167 732. Kymmenen ryhmän ratkaisu: kategoriset muuttujat. Sarakkeet edustavat ryhmiä, rivit kuvaavat ryhmien piirteitä eri taustatekijöittäin.

KLUSTERIT MUUTTUJAT	1 (N=3914 7)	2 (N=917 7)	3 (N=640 0)	4 (N=4228 0)	5 (N=5979 0)	6 (N=4063 6)	7 (N=7500 6)	8 (N=1124 68)	9 (N=2408 18)
<b>Asuinpaikka:</b>	18,7%	16,2%	18,4%	18,9%	12,7%	63,9%	16,5%	14,6%	12,7%
<b>taajama</b>									
<b>maaseutumaiset</b>	20,9%	13,4%	18,4%	19,3%	14,7%	16,2%	56,2%	13,4%	7,8%
<b>Kansalaisuus: suomi</b>	99,6%	98,6%	99,8%	99,6%	98,1%	98,8%	99,5%	97,2%	98,1%
<b>Maassaoloaika: 5-7 vuotta maassa</b>	0,1%	0,4%	0,0%	0,2%	0,3%	0,2%	0,1%	0,2%	0,2%
<b>8-10 vuotta maassa</b>	0,1%	0,5%	0,1%	0,2%	0,4%	0,3%	0,1%	0,2%	0,2%
<b>yli 10 vuotta maassa</b>	0,8%	2,0%	0,5%	0,8%	1,9%	1,2%	0,6%	1,1%	0,8%
<b>Kieli: suomi</b>	96,0%	94,1%	94,2%	95,1%	93,3%	93,1%	92,3%	91,6%	91,8%
<b>ruotsi</b>	0,1%	0,5%	0,1%	0,2%	0,9%	0,3%	0,2%	1,3%	0,9%
<b>Ikä luokiteltuna: alle 24-vuotiaat</b>	0,0%	1,4%	0,0%	0,0%	1,8%	1,0%	0,3%	2,4%	3,0%
<b>24-34 -vuotiaat</b>	0,3%	29,0%	0,0%	0,3%	12,7%	3,9%	1,2%	2,5%	2,0%
<b>35-44 -vuotiaat</b>	0,4%	29,0%	0,1%	0,7%	17,1%	5,6%	1,9%	4,8%	2,3%
<b>45-54 -vuotiaat</b>	0,9%	32,0%	0,3%	0,8%	28,7%	12,6%	4,9%	10,9%	4,5%
<b>yli 65 -vuotiaat</b>	37,8%	0,1%	30,3%	78,5%	0,8%	26,8%	86,1%	0,0%	88,2%
<b>Lasten lukumäärä: Yksi alle 18- vuotias lapsi</b>	0,7%	14,2%	0,4%	0,6%	9,4%	3,3%	0,8%	1,5%	0,6%
<b>2-3 alle 18- vuotiaasta lasta</b>	0,2%	12,4%	0,1%	0,2%	8,9%	2,8%	0,7%	0,3%	0,2%
<b>4 tai enemmän alle 18-vuotiaasta lasta</b>	0,0%	1,2%	0,0%	0,0%	0,9%	0,4%	0,1%	0,0%	0,0%
<b>Asuntokunnan koko: Asuntokunta 1</b>	37,0%	39,0%	24,6%	34,5%	38,2%	21,0%	13,4%	21,0%	28,5%
<b>Asuntokunta 2</b>	54,7%	29,1%	67,5%	57,1%	34,0%	64,9%	75,6%	73,8%	63,5%
<b>Koulutus: erikoisammattitut kinto</b>	0,2%	0,5%	0,6%	0,2%	0,9%	0,4%	0,2%	0,3%	0,2%
<b>liseniaatti tai tohtorin tutkinto</b>	0,2%	0,2%	0,6%	0,2%	0,3%	0,1%	0,4%	0,3%	1,1%
<b>puuttuvat</b>	39,9%	21,0%	26,6%	47,4%	13,1%	74,9%	19,7%	62,3%	47,4%

Taulukko 12. Työkyvyttömyyteen liittyviä etuuksia saanut väestö, N=625 722. Yhdeksän ryhmän ratkaisu: kategoriset muuttujat. Sarakkeet edustavat ryhmiä, rivit kuvaavat ryhmien piirteitä eri taustatekijöittäin.