

Joni Sylvin

# VAPAA-AJAN ELPYMISKOKEMUSTEN JA NIIDEN MUUTOSTEN VAIKUTUS HYVINVOINTIIN

Informaatioteknologian ja viestinnän tiedekunta  
Kandidaattitutkielma  
Toukokuu 2022

# Tiivistelmä

Joni Sylvén: Vapaa-ajan elpymiskokemusten ja niiden muutosten vaikutus hyvinvointiin  
Kandidaattitutkielma  
Tampereen yliopisto  
Tilastollisen data-analyysin tutkinto-ohjelma  
Toukokuu 2022

---

Tämän opinnäytetyön tarkoituksena on tutkia kolmessa eri aikapisteessä kategorisesti mitattujen elpymiskokemusten vaikutusta hyvinvointiin, jota mittaa työssä unen laatuun liittyvä muuttuja. Käytettävän aineiston koko on 616 havaintoyksikköä, eli tutkimukseen osallistunutta henkilöä, joilta kysyttiin sekä sisällä että ulkona olevan mielipaikan kohdalla elpymiskokemuksiin liittyviä kysymyksiä, kuten myöskin unen laatuun liittyviä kysymyksiä.

Sekä vasteena olevasta unimuuttujasta että selittävästä tekijänä olevasta elpymiskokemusmuuttujasta ja niihin liittyvistä erilliskysymyksistä tehtiin summamuuttujat, jotka pääkomponenttianalyysillä tehtiin perusteluina mittaavat samaa asiaa kuin alkuperäiset kysymyssarjat. Sen jälkeen on muodostettiin työn päämenetelmänä olevalla trajektorianalyysillä summaunimuuttujien arvojen perusteella kolmessa eri aikapisteessä 4 eri klusteria, eli sisäisesti mahdollisimman samoilla ominaisuuksilla olevaa aliryhmää, jotta voidaan tutkia tarkemmin erilaisiin klustereihin jakautumisen taustatekijöitä. Kahdessa klusterissa muutoksia ei juurikaan tapahtunut kolmen eri aikapisteen välillä, yhdessä unen heikkenemistä oli lievästi, ja yhdessä klusterissa esiintyi selkeästi unen laadun heikkenemistä siirryttäessä aikapisteissä eteenpäin.

Tämän jälkeen muodostettiin multinomiaalinen logistinen regressiomalli, jossa selitettävänä vastemuuttujana toimii summaunimuuttujien perusteella muodostettu klusteriarvo kullekin havaintoyksikölle. Selittäjäksi valittiin elpymiskokemuksista muodostetut summamuuttujat, joista muodostettiin kategoriset muuttujat kolmeen kategoriaväliin jaettuna elpymissumman arvon perusteella. Lisäksi mukaan otettiin aineistosta muut taustamuuttujat, mukaan lukien ikä, sukupuoli, työ, työtunnit ja koulutus.

Saatujen tulosten perusteella voidaan tehdä johtopäätöksiä, joiden mukaan elpymiskokemuksilla on eri klusteriryhmien vertailussa selkeästi vaikutusta erityisesti viimeisen, eli kolmannen aikapisteen mittauskerran kohdalla, siihen mihin uniongelma-klusteriin henkilö kuuluu. Taustamuuttujista erityisesti työtunneilla ja lisäksi työpaikalla on saatujen tulosten mukaan merkitystä uniongelmien klusteriin kuulumisessa.

Avainsanat: pääkomponenttianalyysi, trajektorianalyysi, klusteri, pitkittäisdata

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -ohjelmalla.

# Sisällys

<b>1</b>	<b>Johdanto</b>	<b>4</b>
<b>2</b>	<b>Tutkimuksen tilastolliset menetelmät</b>	<b>5</b>
2.1	Pääkomponenttianalyysi . . . . .	5
2.2	Trajektorianalyysi . . . . .	5
<b>3</b>	<b>Tutkimusaineisto</b>	<b>7</b>
<b>4</b>	<b>Analyysi valituilla tilastollisilla menetelmillä</b>	<b>8</b>
4.1	Aineiston lukeminen RStudioon ja unimuuttujan analyysi . . . . .	8
4.2	Taustamuuttujien vaikutuksen analyysi multinominaalisen logistisen regressiomallin avulla . . . . .	9
<b>5</b>	<b>Johtopäätökset ja yhteenveto</b>	<b>11</b>
5.1	Päätelmät analyysistä ja mallien sopivuudesta aineistoon . . . . .	11
	<b>Lähteet</b>	<b>12</b>

# 1 Johdanto

Tämän opinnäytetyön tavoitteena on tutkia miten kolmen eri mittauskerran vapaa-ajan elpymiskokemukset ja niiden muutokset ja muu henkilöön liittyvä tausta ennustavat hyvinvointia uniongelmiin kriteerillä. Aiemmin on todettu uniongelmissa ja mielenterveysongelmissa olevan yhteys. Jopa viidelläkymmenellä prosentilla vanhemmista aikuisista on ongelmia unirytmien kanssa, ilmeten vaikeutena nukahtaa tai pysyä unessa. Yhteys uniongelmiin ja mielenterveysongelmiin välillä on kuitenkin aikuisväestössä monimutkainen. (Cable et al. 2016) Tutkimusaineistoksi on saatu professori Kalevi Korpelalta valmisaineisto, jossa on kyselytutkimuksessa mitattu erilaisilla kysymyksillä omaan terveydentilaan, hyvinvointiin ja elpymiskokemuksiin liittyviä asioita, sekä sisällytetty erilaisia taustatietoja tutkimuksiin osallistuneista henkilöistä, mukaanlukien sukupuoli, ikä, koulutus, työ ja työtunnit.

Lopputuloksena halutaan tuottaa ymmärrettäviä ja tulkittavissa olevia tuloksia ja johtopäätöksiä erilaisista tekijöistä, syistä ja taustoista pohjautuen. Luvussa 2 käydään läpi tutkimuksessa tarvittavia tilastollisia menetelmiä, trajektorianalyysiä ja pääkomponenttianalyysiä. Työn päämenetelmänä olevaa trajektorianalyysiä tarkastellaan hieman tarkemmin.

Kolmannessa luvussa esitellään työssä käsiteltävää aineistoa tarkemmin ja neljännessä luvussa toteutetaan varsinainen analyysi käyttämällä esiteltyjä tilastollisia menetelmiä tutkimusaineistoon. Pääkomponentti-, trajektorianalyysi ja multinomiaalinen logistinen regressiomalli ovat siinä vaiheessa keskeisessä roolissa käsiteltävänä olevan pitkittäisdatan ominaisuuksien vuoksi. Luvussa 5 tehdään johtopäätöksiä ja analysoidaan saatuja tuloksia tarkemmin tehdyn analyysin pohjalta ja viimeisenä esitellään työn lähteet.

Tämä tutkielma on kirjoitettu L<sup>A</sup>T<sub>E</sub>X-ohjelmistolla.

## 2 Tutkimuksen tilastolliset menetelmät

### 2.1 Pääkomponenttianalyysi

Pääkomponenttianalyysi on tilastollinen menetelmä, jota käytetään usein tukemaan jotain muuta tilastollista tekniikkaa. Kyseistä menetelmää voidaan myös käyttää perusteluna vaikkapa tiettyjen muuttujien sisällyttämiselle tutkimukseen niiden tilastolliseen malliin sopivuuden ja merkityksellisyyden perusteella. Usein tutkimustilanteessa selittäviä taustamuuttujia voi olla liian monta ja ne voivat lisäksi olla vahvasti keskenään korreloituneita, mikä voi tehdä aineistosta tehtyjen johtopäätösten ja yhteenvetojen tekemisestä hankalampaa.

Menetelmän perusideana on pyrkiä kuvaamaan korreloituneiden taustamuuttujien  $x_1, x_2, \dots, x_n$  eroja kuvaamalla niitä ei-korreloituneilla uusilla muuttujilla  $y_1, y_2, y_3, \dots, y_q$  ( $q < n$ ). (Nummi 2021) Järjestyksessä  $j$ . pääkomponentti on

$$(2.1) \quad y_j = a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jn}x_n = \mathbf{a}'_j \mathbf{x},$$

niin että muuttujan  $y_j$  varianssin arvo on mahdollisimman suuri ehdoilla  $\mathbf{a}'_j \mathbf{a}_j = 1$  ja  $\mathbf{a}'_j \mathbf{a}_i = 0$ , missä kerroinvektorit  $\mathbf{a}_i$  ja  $\mathbf{a}_j$  on estimoitu aineistosta.

### 2.2 Trajektorianalyysi

Työn pääasiallinen tilastollinen tutkimusmenetelmä on trajektorianalyysi, joka on multivariate mixture-mallin erikoistapaus. Analyysi pohjautuu pitkittäisdatan tutkimiseen, jossa eri aikapisteissä on mitattu eri mittauskerroilla tutkittavien muuttujien arvoja, ja näiden mittauskertojen ollessa sarakkeittain matriisimuodossa, voidaan käsiteltävästä datasta laskea klustereittain odotusarvo- ja kovarianssimatriisi. Kyseisessä menetelmässä vastemuuttujan odotusarvo mallinnetaan lineaarisen mallin mukaan, eli kunkin klusterin sisäinen malli on regressiomalli. Odotusarvomatriisi voidaan mallintaa kaavalla

$$(2.2) \quad \boldsymbol{\mu}_k = \mathbf{X}_i \boldsymbol{\theta}_k.$$

Vastaavasti kovarianssimatriisi on muotoa

$$(2.3) \quad \boldsymbol{\Sigma}_k = \sigma_k^2 \mathbf{I}.$$

Klustereiden sisällä havainnot ajatellaan riippumattomiksi, mutta tätä ei voi yleistää koko datan osalta, vaan käytettävään pitkittäisdataan muodostuu erilaisia riippuvuuksia. (Nummi 2021) Trajektorianalyysiä voidaan käyttää mallintamaan yksilöiden erilaisuutta pitkittäisesti. (Nagin, 1999; Nagin, 2005) Sen avulla voidaan käsitellä niin diskreettiä kuin jatkuvaakin dataa.

On tärkeää huomata, että trajektorianalyysillä muodostuneet alijoukot eivät ole kiinteitä, vaan ennemminkin approksimaatioita, joissa tietyllä yksilöllä on tietty todennäköisyys kuulua aliryhmään. (Salonen et al. 2019) Olkoon

$$(2.4) \quad \mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{it})' \quad (t = 1, 2, \dots).$$

tämän työn yhteydessä havaitut unisummamuuttujan arvot yhdeltä henkilöltä  $i$  eri aikapisteessä  $t$ , joita tämän työn sovellustapauksessa on kolme kappaletta. Olkoon  $f_i(\mathbf{y}_i|\mathbf{X}_i)$   $y_i$ :n marginaalitodennäköisyysjakauma, jossa  $\mathbf{X}_i$  kuvaa mahdollisia aika-riippuvaisia kovariaatteja. Tällöin  $f_i(\mathbf{y}_i|\mathbf{X}_i)$  noudattaa  $K$ :n mixturen jakaumaa

$$(2.5) \quad f_i(\mathbf{y}_i|\mathbf{X}_i) = \sum_{k=1}^K \pi_k f_i(\mathbf{y}_i|\mathbf{X}_i, \boldsymbol{\theta}_k), \quad \sum_{k=1}^K \pi_k = 1, \pi_k > 0.$$

Tässä  $\pi_k$  on todennäköisyys kuulua klusteriin  $k$  ja  $f_{ik}(\mathbf{y}_i|\mathbf{X}_i)$  on järjestyksessä  $k$ . tiheysfunktio. (Salonen, et al. 2019) Jotta voidaan löytää parametrien estimaatit, maksimoidaan

$$(2.6) \quad \log L(\boldsymbol{\theta}|\mathbf{y}_1, \dots, \mathbf{y}_n) = \sum_{i=1}^n \log f_i(\mathbf{y}_i)$$

kaikille tuntemattomille parametreille  $\boldsymbol{\theta}$ . Mikäli oletetaan moniulotteinen normaali-jakauma, on  $f_i(\mathbf{y}_i)$  jakauman tiheysfunktio ja  $\boldsymbol{\theta}$  koostuu odotusarvo- ja kovarianssimatriisin parametreista ja suhteellisista osuuksista. (Nummi 2021) Usein monimutkaisissa ja myös trajektorianalyysin tapauksissa voidaan suurimman uskottavuuden estimointiin käyttää EM (Expectation and Maximation)-algoritmia, joka koostuu kahdesta vaiheesta. Vaiheessa E lasketaan logaritmoidun uskottavuusfunktion ehdollinen odotusarvo senhetkisten parametrien voimassaollessa ja senjälkeinen M-vaihe maksimoi odotusarvon mukaisen logaritmitodennäköisyysfunktion. Näitä vaiheita toistetaan niin kauan, että estimaatit alkavat konvertoida ja trajektorianalyysin tapauksessa vaiheessa E lasketaan posterioritodennäköisyys aliryhmään kuulumiselle kaavalla

$$(2.7) \quad \omega_{ik} = \frac{\pi_k f_{ik}(\mathbf{y}_i|\mathbf{X}_i, \boldsymbol{\theta}_k)}{f_i(\mathbf{y}_i|\mathbf{X}_i)}$$

parametriestimaateilla  $\hat{\boldsymbol{\theta}}$ . Estimoidut klusteritodennäköisyydet saadaan laskettua kaavalla

$$(2.8) \quad \hat{\pi}_k = \frac{1}{n} \sum_{i=1}^n \hat{\omega}_{ik}.$$

M-askeleessa käytetään E-askeleessa estimoituja todennäköisyyksiä ns. painoina ja maksimoidaan odotusarvoinen logaritmitodennäköisyys niiden avulla. (Salonen, J. et al. 2019)

### 3 Tutkimusaineisto

Tässä opinnäytetyössä käytetty aineisto on peräisin professori Kalevi Korpelalta saadusta Recovery-aineistosta, jossa on kerätty kolmella eri mittauskerralla tutkimustietoa erilaisilla kysymyksillä vapaa-ajan elpymiskokemuksiin ja niiden muutoksiin liittyen. Mittauskerrat on toteutettu ajallisesti tasavälein ja välin pituus on noin yksi vuosi. (Korpela et al. 2017) Otokoko on 664 ja elpymiskokemusta on mitattu 11 kappaleella väittämiä, joita jokaiselta vastaajalta on kysytty sekä sisä- että ulkotilassa olevan mielifaikan kohdalla. Aineistosta on kuitenkin pudotettu pois puuttuvaa dataa sisältävät havaintoyksiköt, jotta tilastollisten menetelmien käyttö ja muuttujien vertailu onnistuisi, joten tarkastelukohteeksi jää 616 havaintoyksikköä. Puuttuvan datan määrä on suhteellisesti kohtalaisen pieni ja sen oletetaan olevan satunnaista, joten sen ei näin ollen oleteta vaikuttavan tutkimuksen lopputuloksiin.

Aineistossa on jokaiselta tutkimukseen osallistuneelta henkilöltä sukupuoli (nainen arvolla 1, mies 2), ikä, työ (työntekijä (esim. linja-auton kuljettaja, asentaja, siivoaja) 1, alempi toimihenkilö (esim. myyntineuvottelija, assistentti, sairaanhoitaja) 2, ylempi toimihenkilö (esim. opettaja, lääkäri, kehittämisspäällikkö) 3, korkea johtotaso (esim. toimitusjohtaja, osaston johtaja, rehtori) 4), työtunnit viikossa ja koulutustaso (peruskoulu / keskikoulu tai vastaava 1, ammatillinen perustutkinto tai ylioppilastutkinto 2, erikoisammattitutkinto 3, ammatillinen opistoasteen tutkinto 4, alempi korkeakoulututkinto tai AMK-tutkinto 5, ylempi korkeakoulututkinto 6, lisensiaatin tai tohtorin tutkinto 7). (Korpela et al. 2017)

Lisäksi on mitattu kolmessa eri aikapisteessä elpymiskokemuksiin, unen laatuun ja koettuun terveydentilaan liittyviä kategorisella mitta-asteikolla mitattavia kysymyksiä. Tässä työssä selitettävänä vastemuuttujana toimivat unen laatuun liittyvät kysymykset, joista on muodostettu summamuuttuja kullekin havaintoyksikölle erikseen jokaiselle kolmelle aikapisteelle. Myös elpymiskokemuksille on tehty omat summamuuttujansa saman perustelun pätiessä pääkomponenttianalyysien perusteella.

## 4 Analyysi valituilla tilastollisilla menetelmillä

### 4.1 Aineiston lukeminen RStudioon ja unimuuttujan analyysi

Luetaan ensin aineisto RStudioon ja pudotetaan siitä pois puuttuvan tiedon rivit. Alkuperäinen otoskoko on 664, mutta siitä joudutaan vähentämään pois puuttuvaa dataa sisältävät havaintoyksiköt. Tämä tehdään, jotta tilastolliset suureet ja menetelmät voidaan laskea ja toteuttaa datalle. Jäljelle jää 616 havaintoyksikköä. Selitettävänä vastemuuttujana toimivia unimuuttujia on useita jokaisella kolmella eri aikapisteellä aineiston alkuperäisen kysymyssarjan mukaisesti, joten tiedon yksinkertaistamiseksi halutaan muodostaa näistä yksittäisistä unimuuttujista yksi summaunimuuttuja jokaiselle kolmelle aikapisteelle. Poimitaan datasta unimuuttujat aikapisteessä 1 ja lasketaan niiden välinen kovarianssimatriisi, jotta pääkomponenttianalyysiä voidaan hyödyntää tässä yhteydessä.

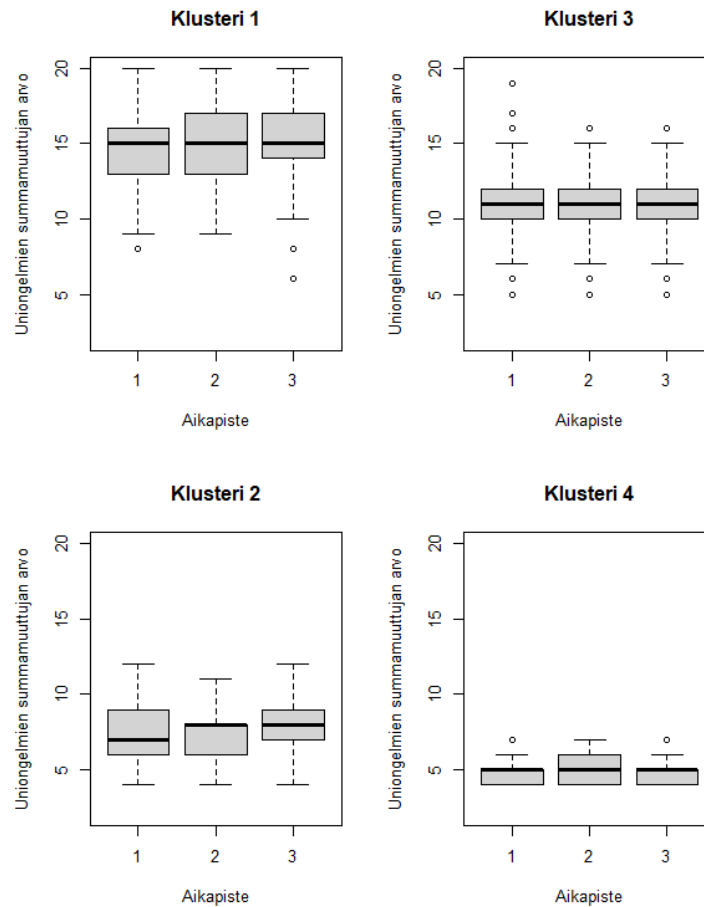
Seuraavaksi lasketaan kovarianssimatriisin ominaisarvot ja pääkomponentit R-ohjelman komennolla *eigen(covsleep)*, jossa *covsleep* on mainittu kovarianssimatriisi. Muodostuneiden pääkomponenttien ensimmäiselle pääkomponentille saadaan unimuuttujittain latauksina -0.43, -0.57, -0.54 ja -0.44, joista voidaan huomata, että painot ovat suhteellisen lähellä toisiaan eri unimuuttujilla. Tämä puoltaa summamuuttujan muodostamisen sopivuutta tilanteeseen, sillä näin voidaan perustella, että summamuuttujalla mitataan edelleen samaa asiaa kuin mitä yksittäiset unikysymysmuuttujat mittaavat kokonaisuudessaan. Lisäksi tulee laskea ensimmäisen pääkomponentin selitysaste, joka lasketaan jakamalla ensimmäinen ominaisarvo kaikkien ominaisarvojen summalla. Näin kaavalla  $3.39/(3.39 + 0.84 + 0.64 + 0.48)$  saadaan selitysasteeksi 0.63, joka on kohtuullisen hyvä selitysaste. Näin ollen summaunimuuttujien käyttö on perusteltua tässä tilanteessa pääkomponenttianalyysin perusteella.

Muodostetaan seuraavaksi unisummamuuttujista matriisi yhdistämällä kaikkien kolmen aikapisteen eri havaintoyksiköiden unimuuttujat ja merkitään jokaiseen niistä tieto aikapisteestä, jotta tieto siitä saadaan pidettyä tallessa. Seuraavaksi summamuuttujavektorit muokataan vektorimuotoon, jotta voidaan hyödyntää R-ohjelmiston kirjastoa FlexMix. Kyseisen kirjaston funktio StepFlexMix suorittaa kerätyille summamuuttujavektoreille aikapisteisiin edelleen jaoteltuna klusterianalyysin trajektorianalyysin menetelmin, jolloin analyysin perusteella voidaan valita esimerkiksi klusterien lukumäärä funktioiden tuottamien informaatiokriteerien perusteella. Informaatiokriteereiden perusteella päädytään valitsemaan neljän eri klusterin malli.

Kuvassa 4.1 nähdään muodostuneet neljä klusteria. Huomataan, että ensimmäisessä niistä uniongelmia näyttää olevan suhteellisen paljon ja lisäksi ne lisääntyvät hieman tarkastelujakson aikana. Toisessa klusterissa uniongelmat selkeästi lisääntyvät ajan myötä ja vertailussa muiden klusterien kanssa uniongelmia on keskiverroksi. Klusterissa kolme muutoksia ei käytännössä näytä tapahtuvan ollenkaan uniongelmien määrän ollessa yleisesti keskimääräinen muihin verrokkiryhmiin nähden, ja



viimeiseen taas ovat kerääntyneet henkilöt, jotka pärjäävät verrokkeihin nähden hyvin ilman unioingelmia.



**Kuva 4.1.** Muodostuneet 4 eri klusteria

## 4.2 Taustamuuttujien vaikutuksen analyysi multinominaalisen logistisen regressiomallin avulla

Edellisellä menetelmällä saatiin jaettua havaintoyksiköt neljään eri klusteriin käytetyn trajektorianalyysin tuloksena. Viedään tieto kunkin havaintoyksikön klusteriarvosta alkuperäiseen havaintomatriisiin ja suoritetaan seuraavana selittävinä muuttujina olevien elpymiskokemusmuuttujien pääkomponenttianalyysi samalla tavalla kuin unimuuttujille tehtiin. Ensimmäisen pääkomponentin painoiksi saadaan -0.35, -0.38, -0.42, -0.42, -0.46 ja -0.42 ja selitysasteeksi  $5.39 / (5.39 + 0.68 + 0.35 + 0.29 + 0.22 + 0.19) = 0.76$ . Näin ollen on perusteltua käyttää myös elpymiskokemusmuuttujien mittarina summattuja elpymiskokemusmuuttujia ja voidaan perustellusti sanoa summamuuttujien mittaavaan samaa asiaa kuin alkuperäisetkin elpymiskokemusmuuttujat mittasivat.

Sovitetaan sen jälkeen multinomiaalinen logistinen regressiomalli aineistoon, niin että selitettävänä muuttujana mallissa on kunkin havaintoyksikön klusteriarvo ja

selittävinä muuttujina henkilön ikä, sukupuoli, työ, koulutustaso ja tehdyt viikkotyötunnit. Lisätään vielä selittäjäksi kunkin aikapisteen elpymissummamuuttujat kukin yksittäisenä selittäjänä niin, että jokainen niistä on vielä tuotu malliin kategorisena luokkamuuttujana kvantiilien perusteella jaettuna kolmeen eri väliin. Tarkoituksena on siis tutkia, onko kyseisillä selittäjämuuttujilla vaikutusta uniongelmiin. Tilastollinen malli saadaan muodostettua R-ohjelmiston VGAM-kirjaston `vglm`-funktiolla, jonka avulla saadaan laskettua kunkin selittävän muuttujan merkitsevyystaso ja kertoimet.

Kyseisen muodostetun mallin referenssitasoksi muodostuu neljäs, eli viimeinen klusteri, mikä tarkoittaa, että muita klustereita verrataan suhteessa neljänteen klusteriin, jossa on selkeästi vähiten uniongelmiä verrattuna muihin. Kun verrataan mallin avulla referenssiklusteria muihin klustereihin, verrattuna ensimmäiseen klusteriin, jossa uniongelmiä on paljon, löydetään merkitsevinä taustamuuttujina mallista tehdyt työtunnit ja viimeisen aikapisteen elpymiskokemukset. Tämän mukaan kyseisiä ryhmiä erottaa tilastollisesti selkeästi työtuntien määrä ja elpymiskokemukset, eli lisääntyvät työtunnit ja toisaalta vähäiset elpymiskokemukset tapaisivat johtaa lisääntyneisiin uniongelmiin. Toiseen klusteriin verrattaessa merkitsevä tekijä on ainoastaan tehtyjen työtuntien määrä. Toisesta klusterista kuvan 4.1 vasemmasta alalaidasta nähdään, että siinä uniongelmissa tapahtuu kolmen aikapisteen välillä nousua, eli suuremmat työtunnit vaikuttaisivat lisäävän uniongelmiä eri aikapisteen välillä. Verrattessa referenssiklusteria kolmanteen klusteriin työ, työtunnit ja elpymiskokemukset kolmannella aikapisteellä nousevat esiin merkitsevinä tekijöinä, joilla on tilastollinen merkitys uniongelmiin. Kolmanteen klusteriin kuuluivat lähinnä henkilöt, joilla oli keskimääräisesti uniongelmiä ilman merkittäviä muutoksia eri aikapisteen välillä. Elpymiskokemusten vähäisyys ja työtuntien isompi määrä korostuvat tilastollisesti kolmannessa klusterissa verrattuna neljänteen. Työn osalta korkeamman statuksen työ näyttää selvästi vähentävän todennäköisyyttä uniongelmiille verrattuna matalamman statuksen työhön.

Tilastollisesti merkitsevien muuttujien osalta tuloksiksi saadaan RStudioon kautta muodostetusta multinomiaalisesta logistisesta regressiomallista seuraavanlaiset riskisuhteet. Yksi lisätyötunti 1.06-kertaistaa riskin kuulua ensimmäiseen uniongelmaisten klusteriin, referenssiklusterin sijaan. Elpymiskokemusten summamuuttujat oli kategorisoitu kolmeen eri kategoriseen muuttujaan ja yhden kategorian parannus elpymiskokemuksissa viimeisessä aikapisteessä pienentää todennäköisyyttä kuulua uniongelmaiseen klusteriin, sen kerroin on 0.45, kun taas ensimmäisen aikapisteen elpymiskokemuksen kerroin on vastaavasti 0.39. Nousevien uniongelmiä klusterin kaksi vertailussa neljänteen klusteriin yksi työtunti lisää nostaa todennäköisyyttä kuulua klusteriin kaksi 1.06-kertaiseksi neljänteen klusteriin nähden.

Keskimääräisten ja muuttumattomien uniongelmiä klusteriin kolme kuulumista verrattuna klusteriin neljä lisäävät tilastollisesti merkittävästi työtunnit 1.06-kertaisesti per lisätyötunti, ja vähensivät työn status 0.50-kertaisesti per työ kategorian yksi parannus kategorisella asteikolla ja kolmannen aikapisteen elpymiskokemuksen kategorian parannus yhdellä 0.46-kertaisesti. Mallin perusteella yksikään taustamuuttujista ikä, sukupuoli ja koulutustaso eivät ole tilastollisesti merkitseviä muiden taustamuuttujien rinnalla.

## 5 Johtopäätökset ja yhteenveto

### 5.1 Päätelmät analyysistä ja mallien sopivuudesta aineistoon

Aineiston ja tehtyjen analyysien perusteella saadaan muodostettua tuloksista johtopäätöksiä ja vastattua tutkimuksen alussa muodostettuihin tutkimusongelmiin. Valitut tilastolliset mallit vaikuttavat suoriutuvan aineiston käsittelystä pääosin hyvin, pääkomponenttianalyysin summamuuttujien käytön perustelua tukeva rooli on tärkeässä asemassa, jotta dataa voitiin muokata yksinkertaisempaan muotoon menettämättä kuitenkaan erillisten kysymysmuuttujien antamaa tietoa. Trajektorianalyysi puolestaan soveltuu erityisesti tämän tutkimuksen kaltaisen pitkittäisdatan käyttöön, kun mittauksia on tehty useammassa eri ajankohdassa. Trajektorianalyysin menetelmästä olisi todennäköisesti ollut vielä enemmän hyötyä mikäli eri aikapisteiden välillä olisi enemmän vaihtelua useammassa klusterissa, mutta toisaalta myös saadunkaltainen tulos on informoiva. Multinomiaalinen logistinen regressiomalli oli erityisen hyödyllinen tutkimuskysymyksen pääasiallisen selittävän muuttujan asemassa olevien elpymiskokemusten tuomisesta tutkittavaksi yhteydestään selitettäviin unioongelmiin, kuten myös muiden lisätaustamuuttujien, kuten iän, sukupuolen, työtuntien, työn ja koulutustason tuomisessa tarkasteluun selittävinä tekijöinä.

Ikä, sukupuoli ja koulutustaso eivät tutkimustulosten ja käytettyjen tilastollisten mallien perusteella nousseet tilastollisesti merkitseviksi tekijöiksi, vaan pääosan unioongelmista näyttävät selittävän erityisesti tehdyt työtunnit, elpymiskokemukset kolmannessa, eli mittauksen viimeisessä aikapisteessä ja osittain myös työtehtävän taso.

# Lähteet

- [1] Cable, N. et al. (2016) *Can sleep disturbance influence changes in mental health status? Longitudinal research evidence from ageing studies in England and Japan. Sleep medicine.* [Online] 30216–221.
- [2] Korpela, K., Nummi, T., Lipiäinen, L., De Bloom, J., Sianoja, M., Pasanen, T., Kinnunen, U. (2017). *Nature exposure predicts well-being trajectory groups among employees across two years.* Journal of Environmental Psychology , doi: 10.1016/j.jenvp.2017.06.002.
- [3] Nagin DS. (1999). *Analyzing developmental trajectories: a semiparametric, group-based approach.* Psychological Methods 4:139–157. DOI: <https://doi.org/10.1037/1082-989X.4.2.139>
- [4] Nagin D. (2005). *Group-Based Modeling of Development.* Cambridge, MA: Harvard University Press
- [5] Nummi, T. (2021). *Multivariate Analysis.* Luentomoniste, Tampereen yliopisto, s. 118.
- [6] Salonen, J. et al. (2019) *Using trajectory analysis to test and illustrate microsimulation outcomes.*