



Aalto-yliopisto
Perustieteiden
korkeakoulu

Analyzing European fertility rates with multiple regression

Aleksi Niemi

13.1.2020

Ohjaaja: *Jukka Kohonen*

Valvoja: *Pauliina Ilmonen*

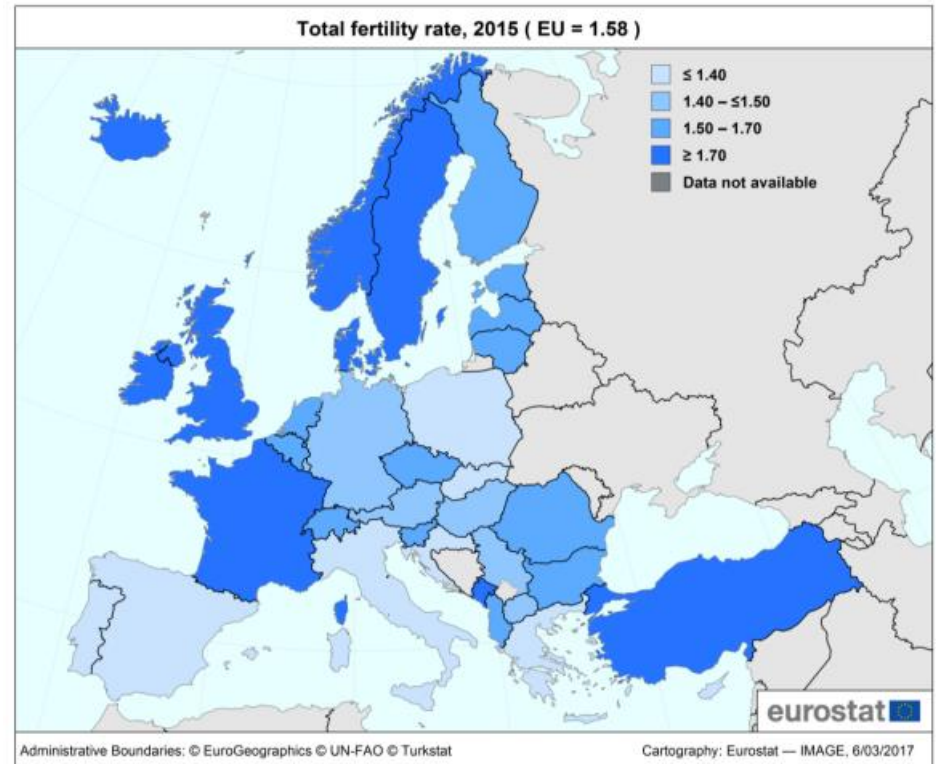
Työn saa tallentaa ja julkistaa Aalto-yliopiston avoimilla verkkosivuilla. Muilta osin kaikki oikeudet pidätetään.

Työn tausta

- Hedelmällisyys, mitattuna naisten odotusarvoisella lapsiluvulla (Total Fertility Rate, TFR), on ollut Euroopassa laskutrendissä viime vuosisadalla
- Hedelmällisyysluvut ovat laskeneet uusiutumisrajan alle
- Myös Suomessa kokonaishedelmällisyys on laskenut ihan viime vuosina rajusti (vuonna 2010 1,87, vuonna 2018 1,41)
- Taustasyistä näkee paljon spekulointia julkisessa keskustelussa
- Syitä kehitykseen kannattaa tarkastella tilastollisesti

Työn tavoite

- Työn tarkoituksena on rakentaa malli, joka ennustaa kokonaishedelmällisyyttä sosioekonomisten muuttujien (esim. bkt, koulutustaso jne.) avulla
- Selvitetään, mitkä muuttujat ovat selitysvoimaisimpia

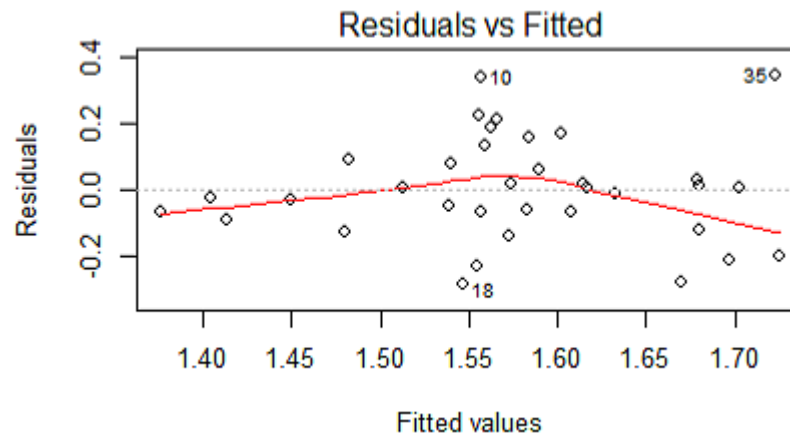


Menetelmät

- Menetelmänä käytetään usean muuttujan regressioanalyysia (multiple regression)
- Mallissa käytettävät selittävät muuttujat valitaan esimerkiksi jo tehtyjen tutkimusten perusteella
- Kokeillaan myös entuudestaan kokeilemattomia selittäjiä
- Regression ajamiseen käytetään RStudio-ohjelmistoa

Menetelmät

- Mallin selitysvoimaa tutkitaan residuaalianalyysillä ja muilla diagnostisilla testeillä



Menetelmät

- Dataa on monelta vuodelta
- Regressio ajetaan tuoreimmalla datalla
- Tuloksia voi validoida ajamalla malli esimerkiksi 10 vuotta vanhalla datalla

Rajaukset ja tietolähteet

- Analyysi on rajattu eurooppalaisiin maihin
- Data hankitaan pääasiassa Eurostatin arkistoista
- Muita lähteitä käytetään tarvittaessa

Aikataulu

- Aihe 11/19
- Esittely 01/20
- Menetelmiin perehtyminen ja regression ajaminen 01/20
- Työn kirjoittaminen 02/20
- Valmiin työn esittely tapaamisessa 03/20 tai viimeistään 04/20