



Aalto-yliopisto  
Perustieteiden  
korkeakoulu

# Bayesian model selection metrics in retail datasets

(valmiin työn esittely)

*Leevi Rönty*

*11.6.2021*

*Ohjaajat: Mikko Ervasti & Paavo Niskala ( Sellforte)*

*Valvoja: Fabricio Oliveira*

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

# Sisältö

- Tausta
- Lähtödata
- Mallit
- Metriikat
- Tulokset
- Yhteenveto

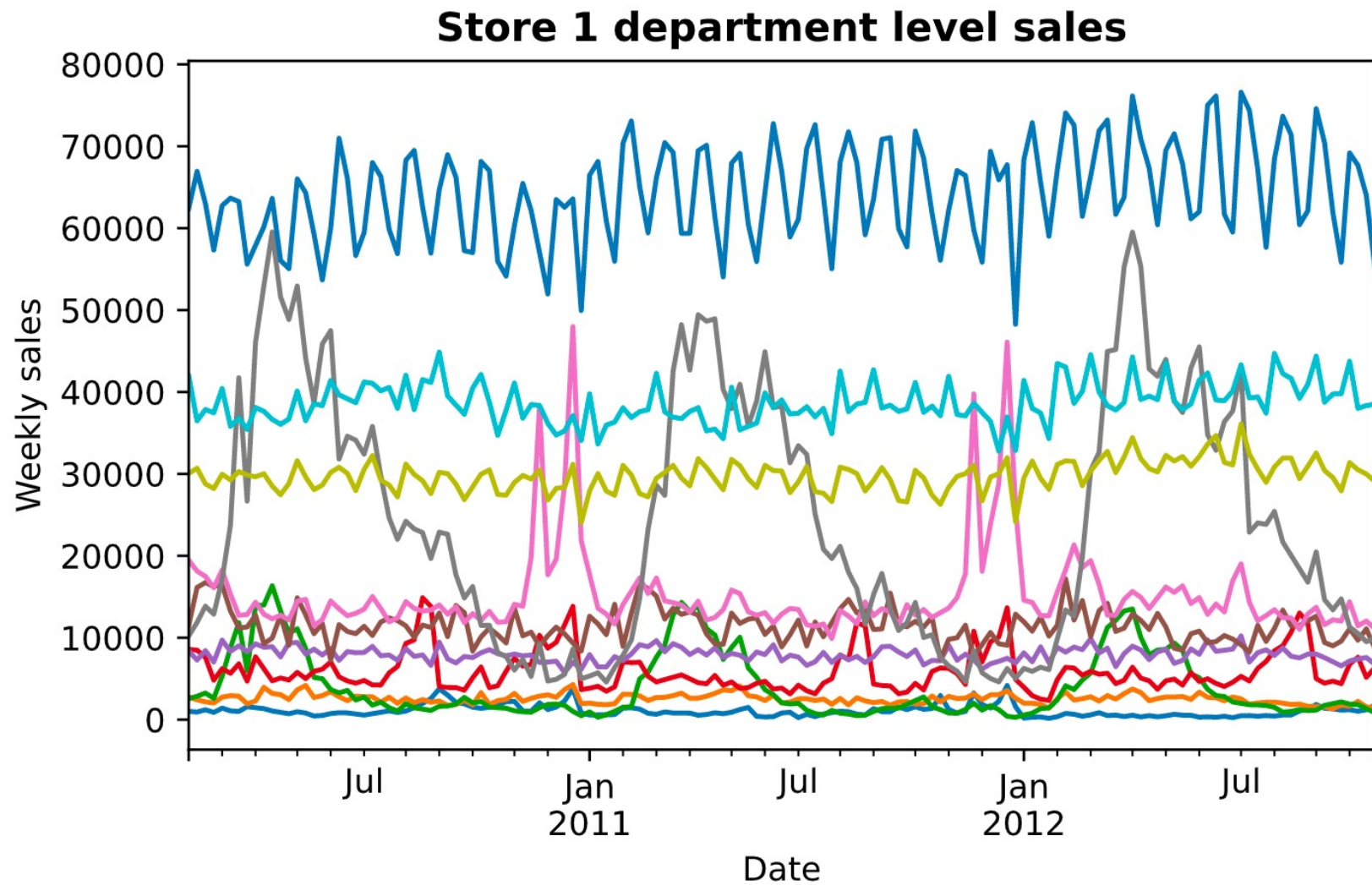
# Tausta

- Bayeslaisten mallien valintaan on useita menetelmiä
- Yksikään metriikka ei kuitenkaan ole täysin ongelmaton kaupallisessa viitekehyksessä
- Kaupallisien mallien kanssa ongelmiksi voivat muodostua mm. laskenta-aika tai datan määrä
- Työn tarkoitus: Tutkia menetelmien käyttökelpoisuutta kaupallisessa ympäristössä. Mitataan evaluointiajat, tutkitaan metriikoiden stabiiliutta, vertaillaan tuloksia eri malleille.

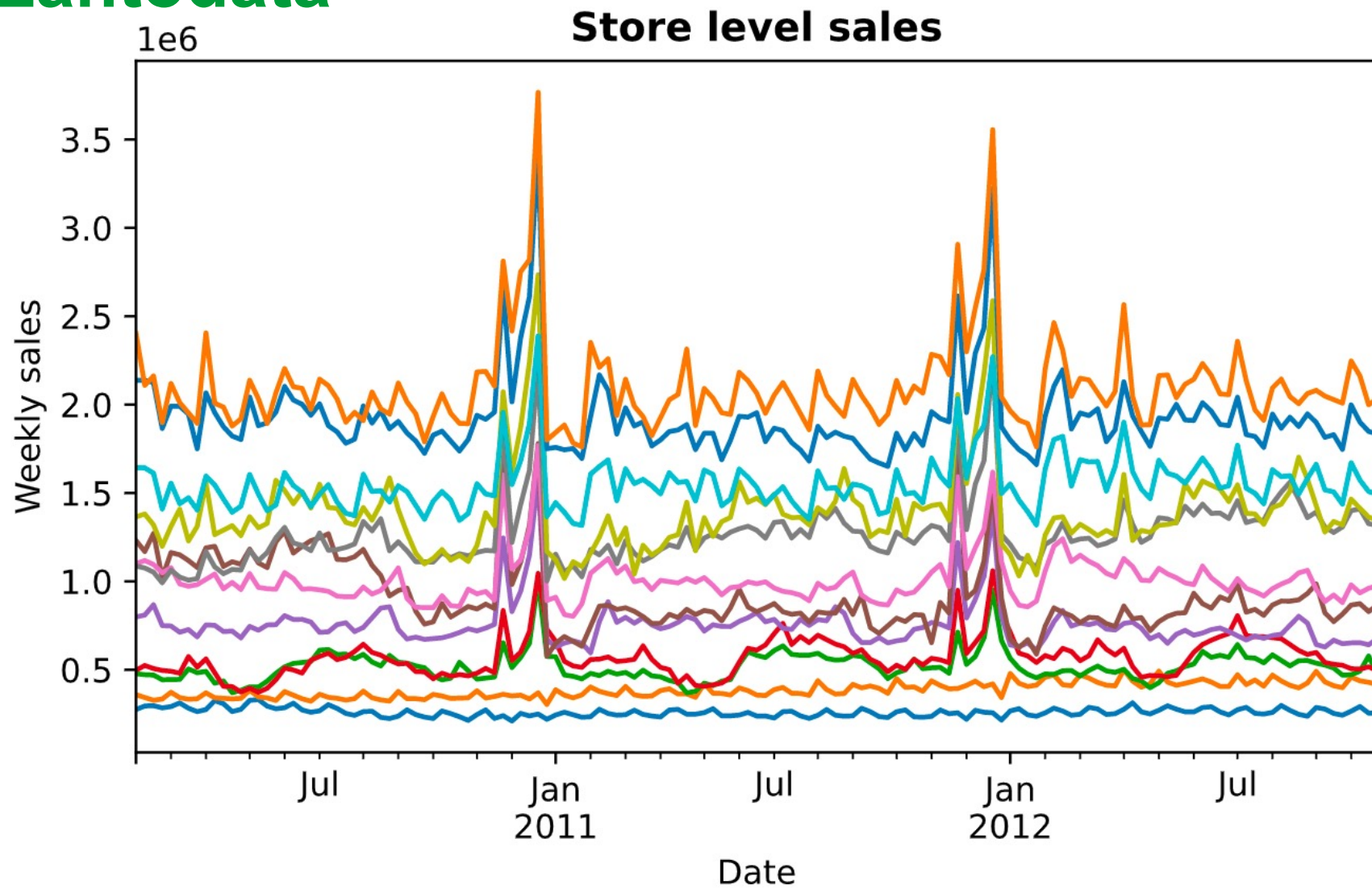
# Lähtödata

- Kaggle open domain datasetti
- Viikoittaista myyntidataa 45 saman ketjun liikkeestä yli kahden vuoden ajalta
- Myynti on eritelty osastokohtaisesti
- Erittelyn tasot: osasto < liike < liiketyyppi
- Lisäksi liikekohtaiset aikasarjat polttoaineen hinnasta, päivän lämpötilasta, työttömyydestä ja kuluttajahintaindeksistä
- Kansalliset erikoispäivät on myös merkattu (esim. Ystävänäpäivä, isot urheilutapahtumat)

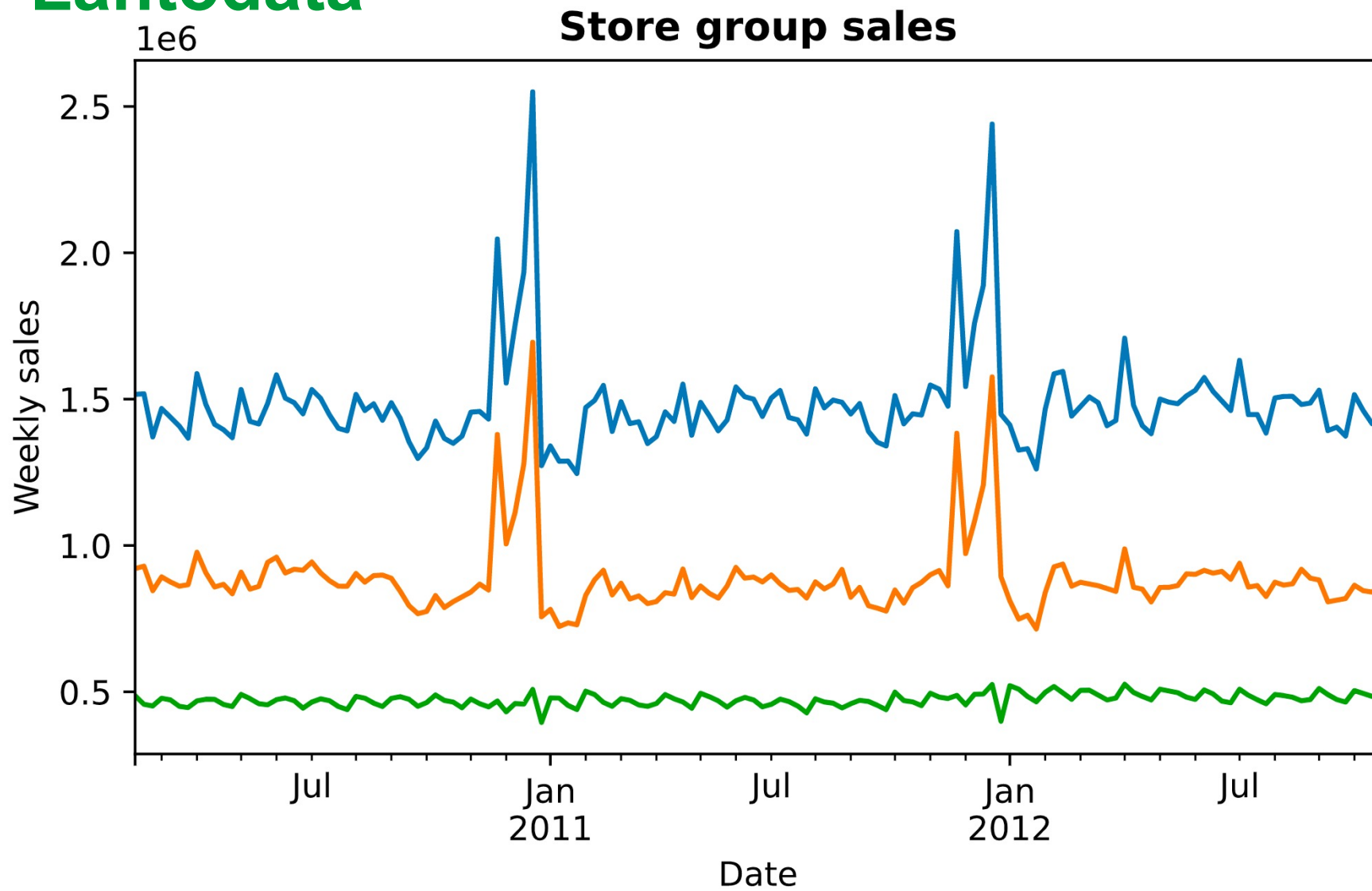
# Lähtödata



# Lähtödata

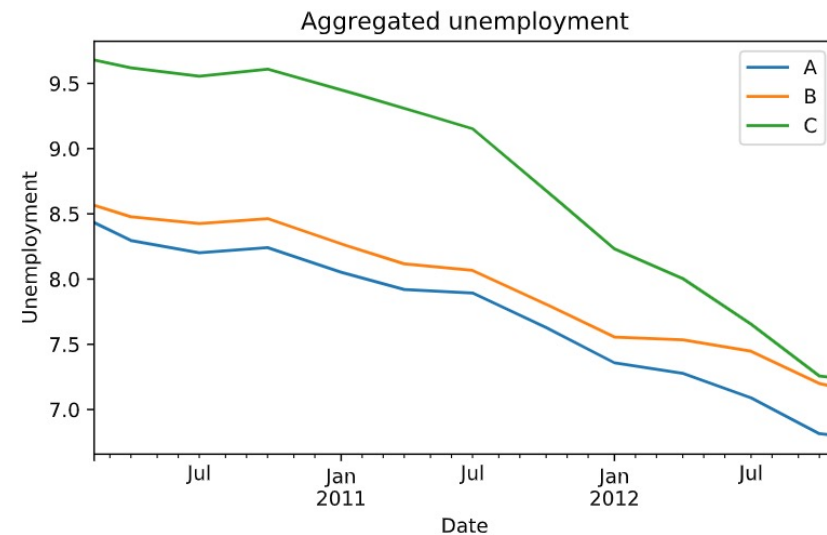
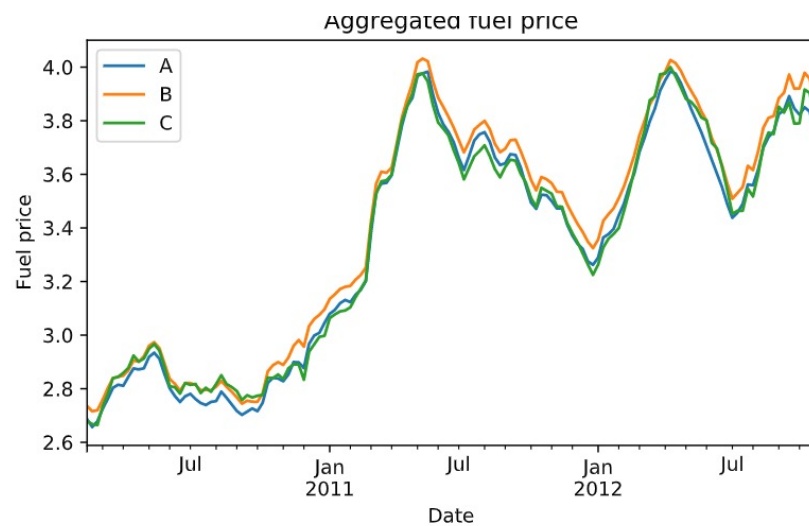
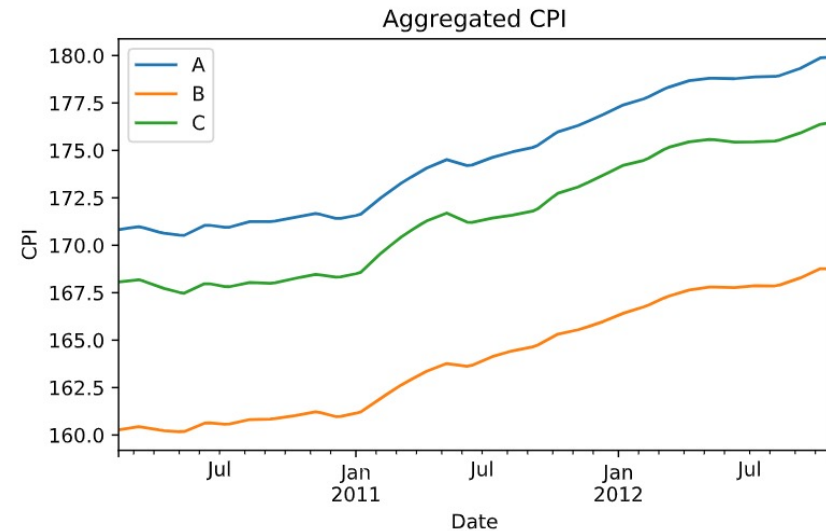
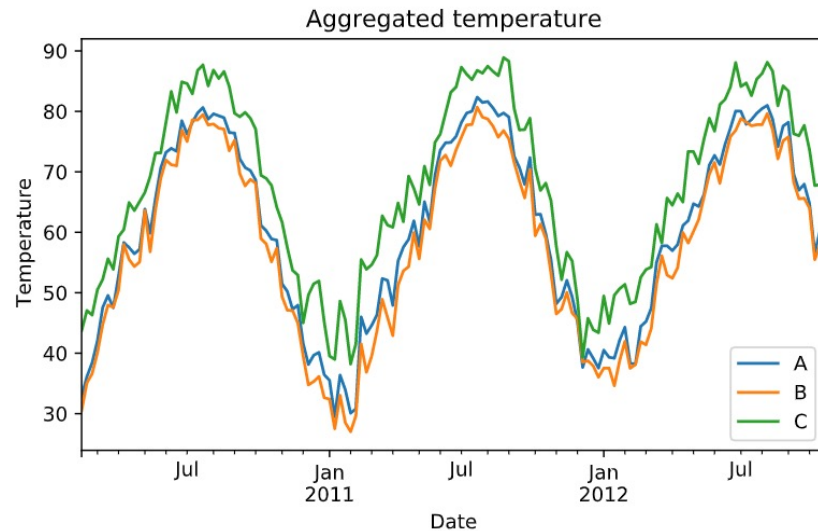


# Lähtödata





# Lähtödata – ulkoiset muuttujat





# Mallit - Prophet

- Facebook Prophet
- Jako kausivaihteluun, trendiin ja ulkoisiin selittäviin tekijöihin
- Kausivaihtelua mallinnetaan äärellisellä Fourier-sarjalla
- Trendinä paloittain lineaarinen trendi, voi käyttää myös logistista trendiä maksimin ja minimin välillä
- Malliin on helppo sisällyttää toimialaosaamista mm. prioreiden tai logististen trendien rajojen avulla.

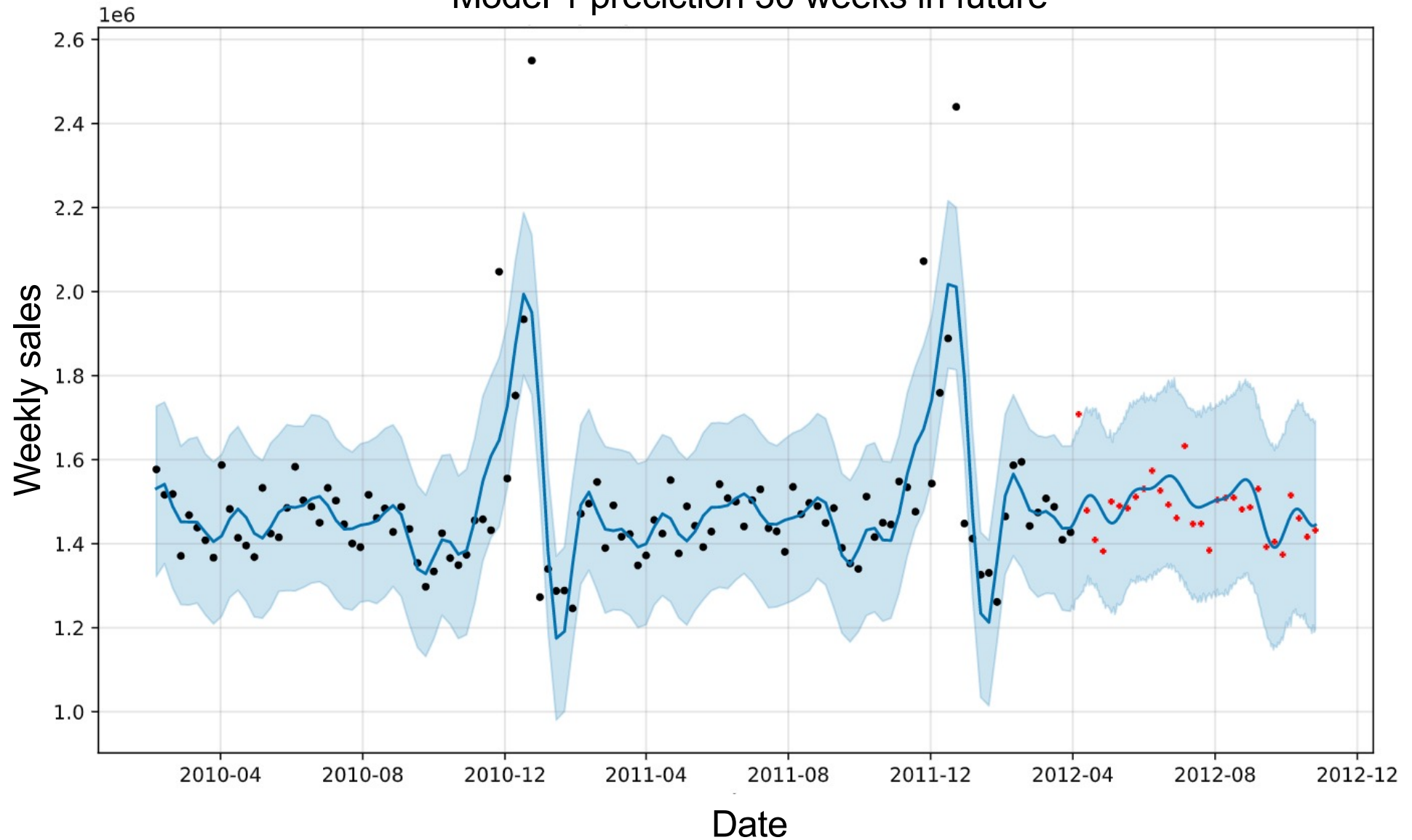
# Mallit

5 eri versiota Prophet-mallista käyttäen eri parametreja ja selittäviä tekijöitä

	M1	M2	M3	M4	M5
Kausivaihtelun kertoimia	10	10	10	20	None
Ulkoiset selittävät muuttujat		X			
Trendin muutospisteet	25	25	25	25	1
y regressorina			X		X

# Mallit – Esimerkki ennusteesta

Model 1 preciction 30 weeks in future



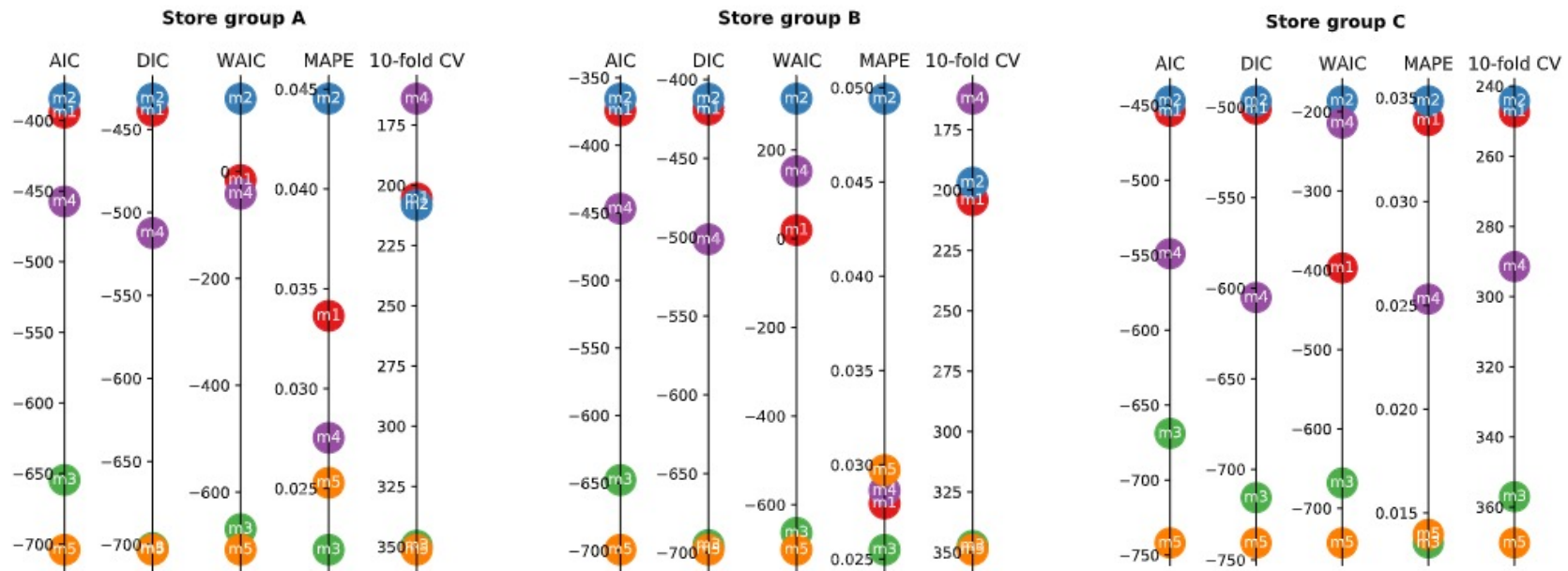
# Metriikat

- Akaike information criterion AIC
- Deviance information criterion DIC
- Widely applicable information criterion WAIC
- MAPE validaatio
- 10-fold cross validation

# Laskutapa

- Informaatiokriteerit lasketaan in-sample –sovitteiden perusteella. Arvio uskottavuudelle, josta vähennetään mallin monimutkaisuudesta riippuva sakko
- MAPE: 15 viikon hold-out setti, lasketaan MAPE koko setille, toistetaan 3 kertaa joka kerta ottaen hold-out setti 7 viikkoa aikaisemmin
- 10-fold CV: jateaan data satunnaisesti 10 osaan, kukin osa toimii vuorollaan hold-out settinä.

# Tulokset – Metriikoiden visualisaatio



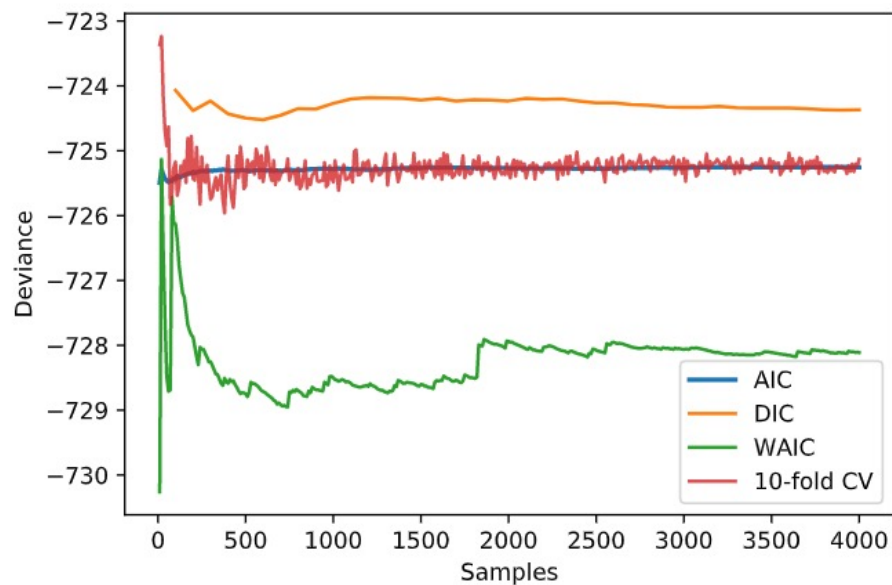
# Tulokset – Metriikoiden evaluointiajat sekunteina

	Initial fitting time	AIC	DIC	WAIC	MAPE	10-fold-CV
M1	32.27	0.00207	8.31	0.00420	108.39	358.46
M2	59.47	0.00326	7.88	0.00412	163.33	491.66
M3	105.55	0.00740	8.06	0.00453	203.68	763.72
M4	65.30	0.00312	7.48	0.00426	151.36	490.07
M5	7.68	0.00376	6.69	0.00453	34.79	83.10

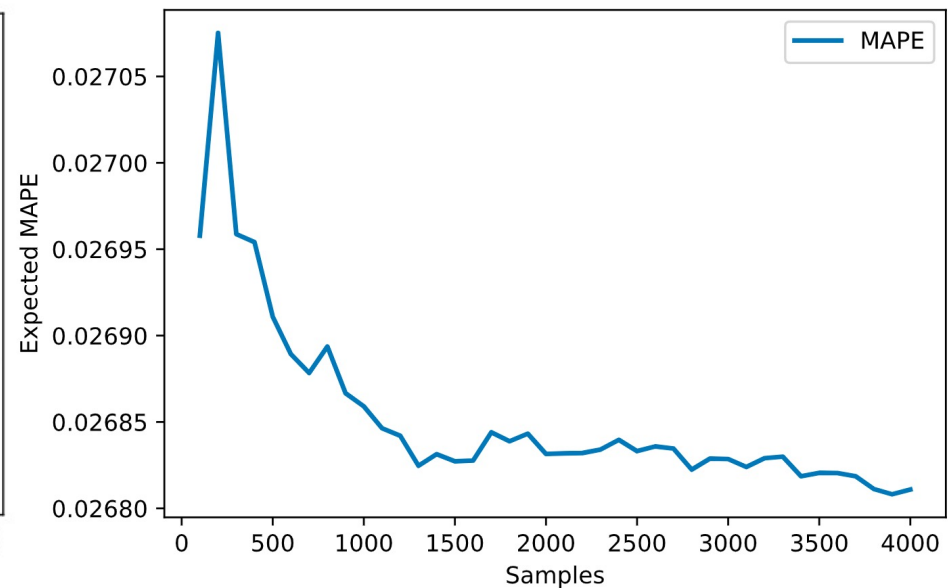


# Tulokset – Metriikoiden vakaus

Metric stability on deviance scale



Metric stability on MAPE scale



# Yhteenveto

- Metriikat arvioivat mallien järjestyksen suurimmaksi osaksi samoin
- Merkittäviä eroja evaluointiajoissa, rajoittaa MAPE:n ja 10-fold CV:n käyttöä etenkin hitaasti sovitettavilla malleilla
- Metriikat käyttäytyivät stabiilisti tarpeeksi suurilla näytemäärillä
- Datan vähyys rajoitti MAPE:n validaatiosovitusten määrää

# Tietolähteet

- Gelman, Andrew and Carlin, John B and Stern, Hal S and Dunson, David B ja Vehtari, Aki ja Rubin, Donald B. *Bayesian Data Analysis*, CRC press, 2013
- Taylor, Sean J ja Letham, Benjamin. *Forecasting at scale*, PeerJ Preprints: 5, syyskuu 2017  
[doi:10.7287/peerj.preprints.3190v2](https://doi.org/10.7287/peerj.preprints.3190v2)
- Vehtari, Aki ja Ojanen, Janne. *A survey of Bayesian predictive methods for model assessment, selection and comparison*, Statist. Surv.:6. 142-228, 2021.