

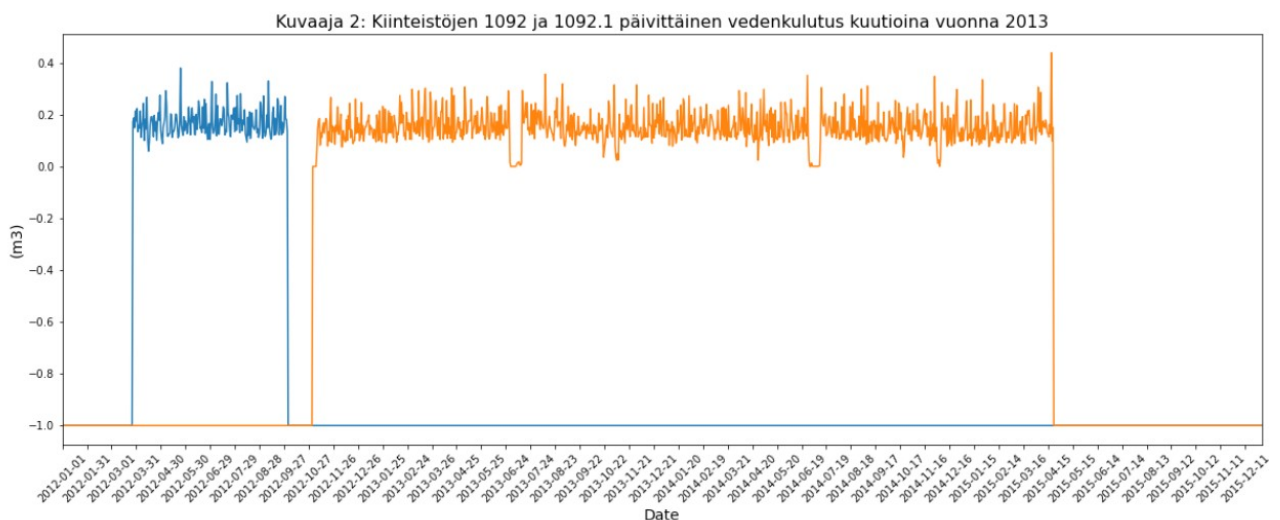
Tehtävä 2 – Datan kuvailu

Tehtävänä on tutkia annetun datan piirteitä. Datan kuvaus on saatavilla osoitteessa <https://datamillnorth.org/dataset/yorkshire-water-daily-customer-meter-data-local-area->, josta käy ilmi mm. datan tuottaja ja ylläpitäjä sekä sen lähteenä toimiva maantieteellinen alue. Lisäksi sivustolta löytyy tieto käytön ehdoista ja se tulee käytettäessä kreditoida. Data itsessään on julkaistu Open Data Commons -lisenssin alla.

Vedenkulutusta koskevassa datassa tarkasteluväli ulottuu vuodelta 2012 vuoteen 2015. Data on saatavilla kiinteistöittäin ja sitä voi tarkastella kiinteistönumeron perusteella. Data on anonymisoitu eikä yksittäistä kiinteistöä ainakaan pitäisi voida yhdistää käytettyyn identifikaattoriin. Osalla kiinteistöistä kiinteistönumero on muotoa 123.1, mikä viitanee omistajan tai vuokralaisen vaihtumiseen. Läheskään kaikissa kiinteistöissä dataa ei ole saatavilla koko ajanjaksolta, vaan saatavuus vaihtelee jonkin verran. Lieneekö tekninen kysymys, ehkä sopimusasia tai riippuvainen mittarin vaihdosta. Yksi datapiste vastaa yhden vuorokauden kulutusta.

Puuttuvan datan kohdalla täytearvona on käytetty negatiivista arvoa, mikä onkin järkevää, koska sille ei kiinteistön veden kulutuksen tapauksessa ole fysikaalista tulkintaa. Näin puuttuva data on välittömästi helppo tunnistaa.

Useimpien kiinteistöjen vedenkäyttö vaikuttaa olevan ympärivuotisesti varsin tasaista. Tyypillinen päiväkulutus näyttäisi vaihtelevan 100 – 300 litran välillä, mikä luonnollisesti selittyy eri talouksien erilaisilla henkilömäärillä. Yksittäisissä tapauksissa erottuu hyvin teräviä kulutuspiikkejä, joiden kohdalla vedenkulutus on moninkertainen. Tällainen yhden päivän mittainen huippu tarkoittanee putkirikkoa tai muuta vahinkotilannetta. Vedenkulutuksen tasaisuus ilman vuodenaikaisvaihtelua herättää epäilyn, että pääsääntöinen vedenkulutus kohdistuu jokapäiväisen elämän toimintoihin, eikä esimerkiksi puutarhan kasteluun.



Kuva 1: Esimerkkikuvaaja erään tarkasteltavan kiinteistön vedenkulutuksesta. Huomaa kiinteistötunnuksen muuttuminen. Nollakulutukset vaikuttavat kesälomilta, mutta muutoin kiinteistön vedenkulutus on tasaista, eikä esimerkiksi vuodenaikaisvaihtelusta näy merkkejä.