

Sään vaikutus Korkeasaaren kävijämääriin

Tekijät:

Jyrki Aho

Janne Andersson

Aki Viitanen

Ohjaava lehtori: Anne-Maritta Talaslahti



Sisällys

1	Johdanto.....	1
2	Empiirinen osa.....	2
2.1	Tavoite ja ongelmat.....	2
2.2	Käytetyt menetelmät ja työvälineet.....	2
2.3	Toteutus- tai työtapakuvaus	3
2.4	Aineisto ja käytetyt analyysit.....	3
2.5	Tulokset ja analyysien yhteenveto.....	4
3	Pohdinta ja oppimisen arviointi	6
3.1	Tutkimuksen yhteenveto.....	6
3.2	Tutkimuksen luotettavuus	6
3.3	Jyrki Ahon oppimisen arviointi	7
3.4	Janne Anderssonin oppimisen arviointi.....	7
3.5	Aki Viitasen oppimisen arviointi	7
	Aineistot.....	8
	Lähteet	8
	Liitteet.....	9
	Liite 1: Korkeasaaren kävijämäärät.....	9
	Liite 2: Ilmatieteenlaitoksen säähavainnot.....	11
	Liite 3: Korkeasaaren sähköposti.....	12
	Liite 4: Korkeasaaren sähköpostin liitetiedosto	14
	Liite 5: Kooste Regressiossa käytetystä datasta	16
	Liite 6: Päättäntäpuun muodostamisessa käytetty data	17
	Liite 7: Päättöspuu.....	18
	Liite 8: Regressiomalli ja logistinen malli	19
	Liite 9: Logistisen mallit	20

1 Johdanto

Sää vaikuttaa olennaisesti ulkoilmatapahtumien viihtyvyyteen ja juuri Korkeasaaren eläintarhassa kävijöiden on liikuttava ulkoilmassa. Vuonna 1889 perustettu Korkeasaaren eläintarha valikoitu kohteeksemme, koska se on Helsingin suosituimpia nähtävyyksiä ja eräs maailman vanhimmista eläintarhoista. Korkeasaari on nykyään auki vuoden jokaisena päivänä, ja sinne pääsee autolla Mustikkamaalta sekä vesibussilla Kauppatorilta ja Hakaniemestä. Mikäli Korkeasaari voisi ennustaa kävijämäärät suhteellisen tarkasti, niin he voisivat saavuttaa säästöjä valmistamalla oikean määrän ruokaa ravintoloissansa, sekä kohdistamalla työntekijöittensä työpanokset oikeisiin kohteisiin. Korkeasaaresta saamamme vastaus loi myös uskoa sille, että voimme löytää yhteyden sään ja kävijämäärien välillä.

Tässä raportissa on tarkasteltu sään vaikutusta Korkeasaaren eläintarhan kävijämääriin viiden vuoden ajalta (2015–2019). Ulkohuvipuistot sekä ulkoilmatapahtumat ovat hyvin sääherkkiä, joten säähän varautuminen on liiketoiminnan kannalta perusteltua. Tämä raportin tarkoituksena on faktatietoon perustuen osoittaa sään vaikutus Korkeasaaren kävijämääriin ja näin ollen tarjota työväline Korkeasaaren johdolle henkilökunnan ja palveluiden resursointiin sekä tukemaan eläintarhan markkinointia.

Käytössämme on avoimista tietokannoista saadut Ilmatieteenlaitoksen päiväkohtaiset säätiedot, sekä Korkeasaaren päiväkohtaiset kävijämäärät. Korkeasaaren ystävällinen henkilökunta puolestaan luovutti heidän aukioloaikansa kyseiseltä ajanjaksolta. Yhdistämällä säädatan, aukioloajat ja kävijämäärät samaan taulukkoon, toivomme analysointityökalujen antavan paremman käsityksen siitä, mitkä sään osatekijät vaikuttavat kävijämääriin. Raportin tarkoituksena on muodostaa tilastoihin perustuvat päätelmät sään vaikutuksesta kävijämääriin viiden vuoden seurantajaksolta.

Teoriamme on, että lämpimät ja pilvettömät päivät houkuttelevat enemmän väkeä Korkeasaareen kuin sateiset säät. Toisaalta pohdimme myös sitä, kuinka runsasta veden tulon on oltava, jotta se vaikuttaisi olennaisesti kävijämääriin. Pohdimme myös miten peräkkäiset sadesäät vaikuttavat kävijämääriin. Teimme myös oletuksen, että kovat tuulet ja erittäin kylmät talvipäivät karkottavat myös kävijöitä. Ryhmämme kykeni eliminoimaan tuulen suunnan ja ilmanpaineen vaikutuksen kävijämääriin. Eräs mielenkiintoinen kysymys oli, kuinka iso ero kävijämäärissä syntyy näiden sääilmiöiden aikana. Koska kenelläkään ryhmällämme ei ole kokemusta ulkoilmatapahtumista, niin tämä kysymys jäi avoimeksi.

2 Empiirinen osa

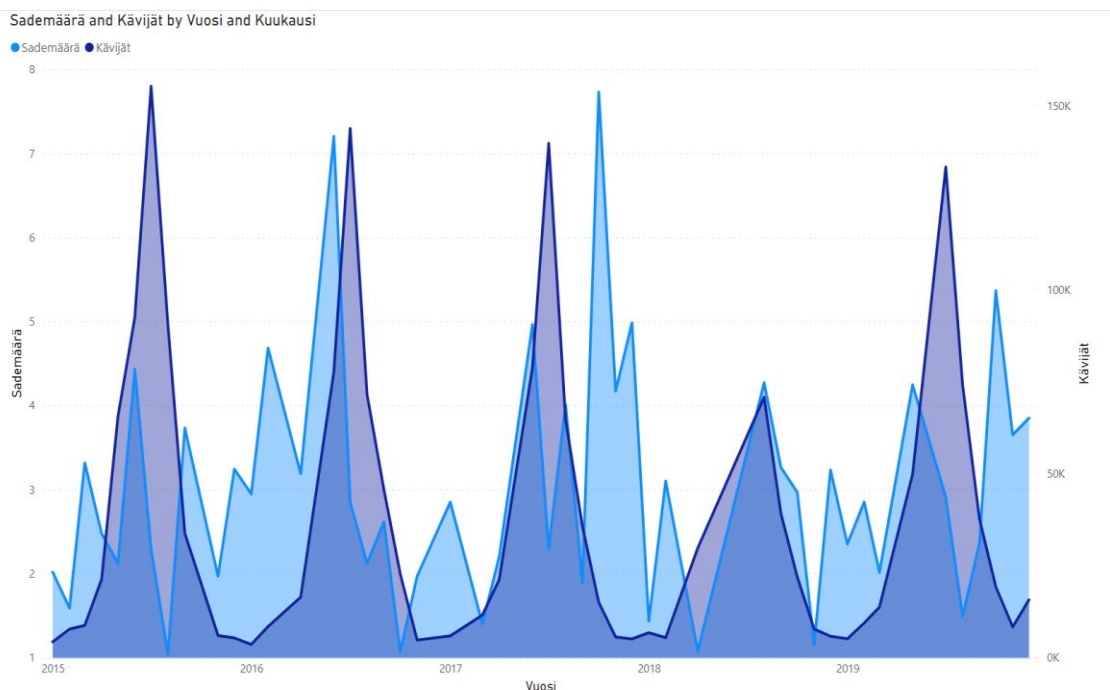
Tässä luvussa kerrotaan projektimme tavoitteista, käytetyistä menetelmistä ja aineistosta saaduista tuloksista.

2.1 Tavoite ja ongelmat

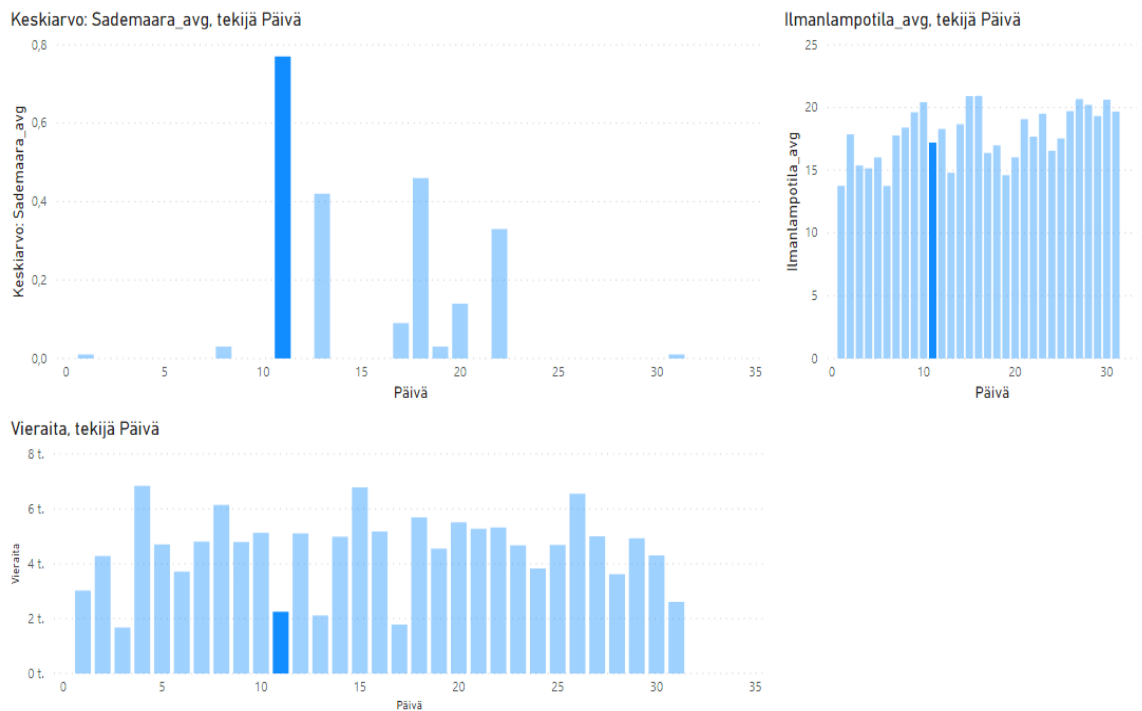
Tässä raportissa selvitetään sään ja kävijämäärien yhteyttä toisiinsa. Tavoitteena on tarjota työkalu Korkeasaaren eläintarhalle resurssisuunnitteluun ja toiminnan kehittämiseen. Sää ei ole pelkästään aurinkoa tai sadetta, vaan se on hyvin moninainen ilmiö ja tarkoituksemme on selvittää, mitkä osatekijät vaikuttavat kävijämääriin. Toimivien ennustemallien avulla voidaan ennustaa esimerkiksi, kuinka paljon ruokaa varataan tarjolle tai kuinka monta kassaa on pidettävä auki. Tarkoituksemme on yrittää muodostaa aineistosta myös päätäntäpuu, jota johtajisto voisi käyttää työkaluna tehostaessaan toimintoja.

2.2 Käytetyt menetelmät ja työvälineet

Työvälineinä ryhmämme käytti Exceliä datan prosessointiin ja luokitteluun. Tavallista tekstieditoria käytettiin apuna, jotta data saatiin muokattua Wekalle sopivaan muotoon. Analysoinnissa ja päätäntäpuun tekemisessä käytettiin apuna Wekan data mining -työkaluja. Aineiston visualisoinnissa ryhmämme käytti PowerBin tarjoamia työkaluja.



Kuva 1. PowerBi:hin ladattiin Excel taulukko, jossa oli sademäärän keskiarvo ja tarkka kävijämäärä jokaiselta päiväältä viiden vuoden ajalta. Vaalean siniset huiput esittävät sademäärää ja tummemman siniset kävijämäärää. Kävijämäärän piikki osuu kesä-elokuulle.



Kuva 2. esittää heinäkuisen tiistaipäivän vuonna 2017, joka on ollut kuukauden sateisin vuorokausi ja kävijämäärä on ollut kolmanneksi pienin. Datan mukaan maanantai on yleensä hiljaisin päivä, kun taas lauantai ja sunnuntai vilkkaimpia. Muuten kävijämäärät jakautuvat melko tasaisesti muille päiville.

2.3 Toteutus- tai työtapakuvaus

Jaoimme ryhmätyömme kolmeen eri osaan, joka osin johtui nykyisestä Covid –tilanteesta ja toinen syy oli taata tutkimuksen luotettavuus. Mikäli kaikki kolme tutkijaa löytäisivät samat selittävät tekijät ilmiön taustalla, niin tutkimustamme voitaisiin pitää luotettavampana. Teimme ryhmässä kuitenkin sen päätöksen, että kaikki tutkivat aineistoa eri analysointimenetelmin. Koska Jyrki Aholla oli eniten kokemusta aineiston käsittelystä ja analysoinnista, niin ryhmä päätti nimittää hänet projektipäälliköksi. Projektiryhmän muina jäseninä toimivat Janne Andersson ja Aki Viitanen.

2.4 Aineisto ja käytetyt analyysit

Ilmatieteenlaitoksen säähavaintojen ja Korkeasaaren kävijämäärät löytyivät avoimista lähteistä, mutta Korkeasaaren aukioloajat eivät olleet julkista tietoa, joten pyysimme nämä tiedot Korkeasaaresta. Aineisto, alkutarkastelussa havaitsimme suoraan, ettei ilmanpaineella ja pilvisyydellä ollut mitään vaikutusta kävijämääriin. Käytimme apuna myös Youtubesta löytyviä opetusvideoita Wekan käytöstä analysoinnissa.

Regressioanalyysi ja varianssianalyysi ovat erittäin tärkeitä tilastollisten mallien työkaluja. Ne perustuvat matemaattisiin työkaluihin, joiden avulla luodaan yksinkertaistettu ennustemalli datasta. Yleisin mallin hyvyyden mittaustapana käytetään mallin jäännösvirheen neliön laskentaa. Mitä pienempi se on, niin sitä tarkemmin mallin odotetaan kuvaavan dataa (Rasch ja Schott, 2018). Toisin ilmaistuna, mikäli tietokone on luonut kaksi eri mallia datasta ja toisen laskennallinen tarkkuus on parempi, niin tietokone valitsee automaattisesti tämän tarkemman mallin jatkotarkasteluun.

Päätöspuut ovat tapa kuvata päätöksentekoa visuaalisesti puurakenteen avulla. Päätöspuukaavio kuvaa päätöksentekoon vaikuttavia tietoja ja ennustettuja lopputuloksia, joka siten toimii päätöksenteon apuvälineenä. Päätöspuun käyttö vaatii, että ennustettava data luokitellaan ennalta sovittujen sääntöjen mukaisesti. Päätöspuun toimintaperiaate on muodostaa datasta optimoituja päätäntäsolmuja ja ennusteita, joissa kuljetaan uuden tiedon perusteella ja tarkastellaan puun lehdessä sijaitsevaa ennustetta. Koneoppimisessa päätöspuu toimii ennakoivana mallina, joka luodaan käytössä olevasta datasta. Tässä tehtävässä päätöspuu on mallinnettu avoimen Wekan avoimeen lähdekoodiin pohjautuvan tiedonlouhinta työkalun avulla, joka kykenee analysoimaan ja visualisoimaan tiedot.

2.5 Tulokset ja analyysien yhteenveto

Koska Korkeasaaren aukioloajat vaihtelevat vuodenaikojen mukaan, niin päätimme eliminoida aukioloaikojen vaikutuksen aineistosta. Otimme huomioon myös sen, että Korkeasaaren lipunmyynti menee tuntia aikaisemmin kiinni kuin Korkeasaari sulkeutuu. Tämän perusteella muodostimme parametrin vierailijoita per tunti. Kun tarkastelimme tätä parametria ja sen luonnollista logaritmia aikajanalla, niin havaitsimme selkeän trendin, joka noudatti sinikäyrää. Tämän perusteella päätelimme ajalla olevan merkittävä rooli selittäessä Korkeasaaren kävijämääriä. Liitteessä 8 esitetty sinikäyrä mallinsi parhaiten ajan vaikutusta kävijämäärien logaritmiin, jonka perusteella laskimme kävijähuipun asettuvan 15.7.–25.7. väliseen aikaan. Kumpulan sääasema ei tallentanut lämpötiloja 2.11.2017 – 6.6.2018 väliseltä ajalta, jotka siten korvasimme Kaisaniemen sääaseman havaintoarvoilla. Stolwijk, Straatman ja Zielhuis (1999) artikkelin pohjalta loimme regressiomallin, joka on nähtävillä liitteessä 8. Regressiomallin avulla havaittiin, että ajalla ja pienelläkin sademäärällä on suuri vaikutus kävijämäärään. Ilman lämpötila nostaa hieman kävijämääriä ja tuuli vähentää lähes yhtä voimakkaasti kävijämääriä. Ilmankosteus vaikuttaa puolestaan jonkin verran kävijämääriin. Weka tuottaman regressiomallin korrelaatiokerroin on 0,7436 ja sen suhteellisen absoluuttisen virhe on 54,189 %. Mikäli mallista jätettäisiin aika

selittävät tekijät pois, niin tällöin korrelaatiokerroin on 0,6803 ja suhteellinen absoluuttinen virhe on 62,5549 %.

Jotta aikatrendi ei vaikuttaisi liikaa aineiston käsittelyyn, niin luokittelimme asiakasmäärät kuukausittain. Oletimme asiakasmäärien noudattavan normaalijakaumaa $V \sim N(\mu_{kk}, \sigma_{kk})$, jolloin asiakasmäärät voidaan luokitella seuraavasti.

Selite	Vierasmäärä kuukausitasolla (V)	Todennäköisyys
Erittäin paljon	$\mu_{kk} + 0,8416 * \sigma_{kk} \leq V$	20 %
Paljon	$\mu_{kk} + 0,25335 * \sigma_{kk} \leq V < \mu_{kk} + 0,8416 * \sigma_{kk}$	20 %
Normaali	$\mu_{kk} - 0,25335 * \sigma_{kk} \leq V < \mu_{kk} + 0,25335 * \sigma_{kk}$	20 %
Vähäinen	$\mu_{kk} - 0,8416 * \sigma_{kk} \leq V < \mu_{kk} - 0,25335 * \sigma_{kk}$	20 %
Erittäin vähän	$V < \mu_{kk} - 0,8416 * \sigma_{kk}$	20 %

Stolwijk, Straatman ja Zielhuis (1999) artikkelin vuoksi tarkastelimme Wekan avulla logistisen malli ennustusta, joka onnistui luokittelemaan oikein 47,5 % tapauksista. Wekan tuottaman logistisen mallin tulokset ovat nähtävissä liitteessä 9.

Liitteessä 7 esitettyyn päätöspuuhun valitsimme datasta mukaan viikonpäivän, lämpötilan, sademäärän tuulennopeuden ja kävijämäärän. Myös tällöin kävijämäärät luokiteltiin aiemmin esitellyn taulukon mukaisiin luokkiin. Muut sään osatekijät jätimme pois, koska niillä ei vaikuttanut olevan vaikutusta kävijämääriin tai ne eivät parantaneet mallin tuloksia. Wekan tuottama päätöspuu ennustaa 46,62 prosenttisesti oikein, joten täysin toimivaa päätöspuuta emme onnistunut Wekalla rakentamaan. Toisaalta mikäli päätöspuun tuottama ennuste olisi ollut satunnainen, niin se olisi ennustanut noin 20 % oikein. Päätöspuusta on kuitenkin nähtävissä trendi, että viikonpäivällä, sateella ja ilmanlämpötilalla on suuri vaikutus asiakasmääriin.

3 Pohdinta ja oppimisen arviointi

Seuraavissa osioissa tarkastellaan tutkimuksen tuloksia ja luotettavuutta. Kappaleen lopussa kukin ryhmän jäsen pohtii omaa oppimistaan tällä kurssilla.

3.1 Tutkimuksen yhteenveto

Aineistoa tutkittaessa havaittiin, että ilmanpaine noudattaa normaalijakaumaa ja pilvien määrä tasaista jakaumaa suhteessa vierailijoiden määriin. Tämän perusteella eliminoimme nämä muuttujat pois tarkasteluistamme. Eliminoimme myös lumen määrän pois, koska lunta näytti olevan runsaasti ainoastaan alkuvuodesta, eikä se näyttänyt vaikuttavan vierasmääriin.

Aineistoa tarkastelemalla havaitsimme ajalla, viikonpäivällä ja sateella olevan suurin vaikutus Korkeasaaren kävijöiden lukumääriin. Toisaalta ajan merkitys on ymmärrettävää, koska kesäisin pidetään pitkiä kesälomia ja matkustellaan enemmän. Suomessa liikkuu kesäisin myös hyvin paljon ulkomaalaisia turisteja, mikä sekin saattaa osaltaan selittää korkeat kävijämäärät. Lauantaisin kävi kaikista eniten vierailijoita Korkeasaarella ja hyvänä kakkosena oli sunnuntai. Maanantaisin kävi kaikista vähiten vierailijoita. Sademäärällä oli suurempi vaikutus kävijämääriin kuin projektin alussa oletimme. Jo pienetkin sademäärät alensivat kävijöiden määrää huomattavasti. Ilman lämpötila kasvatti kävijämääriä, kun taas tuuli ja suhteellinen kosteus alensivat sitä jonkin verran.

3.2 Tutkimuksen luotettavuus

Tutkimuksemme auttoi luomaan kattavan kuvan sään vaikutuksesta Korkeasaaren kävijöiden lukumääriin. Tutkimuksemme tiedot on kerätty avoimista lähteistä ja kolme eri henkilöä tutki aineistoa tietämättä toistensa tuloksista. Koska tuloksemme tuottivat samankaltaiset tulokset, niin tutkimuksemme tulosta voidaan pitää erittäin luotettavana. Tarkastelimme myös peräkkäisten pouta, sekä epävakaisten ja sadepäivien vaikutusta, mutta emme havainneet näillä olevan vaikutusta kävijämääriin. Tutkimuksemme aineisto ei kyennyt täysin selittämään Korkeasaaren kävijämääriä. Mahdolliset ennakkoon sovitut ryhmäopastukset, kuten koululaisryhmien käynnit, sekä mahdolliset mainoskampanjat saattoivat vaikuttaa tutkimuksemme luotettavuuteen. Koska nämä kyseiset tiedot eivät olleet avoimia, niin voimme vain arvailla näiden vaikutusta datan luotettavuuteen. Mahdollinen data turistien lukumääristä pääkaupunkiseudulla saattaisi auttaa osaltaan selittämään Korkeasaaren kävijämääriä. Mahdolliset jatkotutkimukset katsotaan olevan tarpeellisia, jotta saataisiin selville, onko ajalla niin tärkeä rooli selitettäessä kävijämääriä vai vaikuttaako lämpötila ja sademäärät kaikista eniten kävijöiden lukumääriin.

3.3 Jyrki Ahon oppimisen arviointi

Jyrki Aholle aineiston käsittely ja tilastoanalyysit olivat tuttuja jo aikaisemman koulutuksen perusteella. Siten aineiston alkukäsittely, prosessointi ja tutkiminen ei ollut uutta. Toisaalta Weka ja PowerBi olivat ohjelmina uusi tuttavuus, mikä vaati hieman perehtymistä asiaan. Etenkin PowerBi järjestelmään kirjautuminen tuotti aluksi hieman ongelmia ja se vaikutti hyvin mielenkiintoiselta työkalulta muodostaa visuaalista dataa aineistosta. Aikomukseni on perehtyä erittäin huolellisesti sen toimintaan. Wekaan olisin kaivannut hieman parempaa ohjeistusta, koska kyseistä ohjelmia raapaistiin kurssilla hyvin pintapuolisesti. Eli mitä muita asioita kyseisellä ohjelmalla olisi voinut tehdä, mikä olisi edesauttanut aineiston prosessointia. Kurssi oli ihan mukava, mutta ikävä kyllä tunsin suurimman osan kurssin aineistosta entuudestaan.

3.4 Janne Anderssonin oppimisen arviointi

Minulla ei ollut mitään ennakko-odotuksia kurssista, mutta pian selvisi, että kyse ei ole lähipuutojutusta. Kurssin tehtävät olivat haastavia ja osittain myös hieman vaikeasti ymmärrettäviä. Käytettäviin työkaluihin ei ollut riittävästi ohjausta tarjolla, mutta onneksi ryhmän tuella Wekan ja PowerBi:n käyttö onnistui kohtalaisesti. Käytetyistä ohjelmista PowerBi oli mukavampi käyttää, koska sen avulla pystyi tekemään hienoja visualisointeja datasta. Lisäksi Microsoftille tyypilliseen tapaan, ohjelma oli selkeä, melko helposti opeteltava ja se noudatti Office työkalujen toiminnallisuutta. Toistaiseksi en usko, että tulen käyttämään kumpaakaan edellä mainittua ohjelmaa datan analysoimiseen tai visualisointiin, mutta ilman kurssia en olisi tiennyt niistä mitään. Aina on mukavaa oppia jotain uutta, vaikka opeteltavaa asiaa ei sisäistäisi täydellisesti. Aki ja Jyrki ovat todella aktiivisia ja tiedonjanoisia opiskelutovereita, joiden kanssa oli ilo tehdä töitä. Jyrkin taustaa tilastotieteilijänä ei voi liiaksi suitsuttaa.

3.5 Aki Viitasen oppimisen arviointi

Kurssi oli tähänastisen ammattikorkeakoulu uran haastavin. Aihe on erittäin mielenkiintoinen ja kurssista saikin hyvän peruskäsityksen mitä kaikkea tiedon louhimisella ja analysoinnilla pystyy tekemään. Weka ja Power Bi olivat molemmat ihan uusia tuttavuuksia ja luulenkin, että Power BI tulee olemaan jatkossakin käytössä. Myös minä olisin kaivannut Wekan käyttöön lisäohjeita, jotta siitä olisi saanut enemmän irti. En ole niin analyyttinen ihminen, että jatkossa minua kiinnostaisi analysoida näin isoja datamääriä, mutta kurssin opit tulevat varmasti käyttöön jossain määrin. Minulla on oma verkkokauppa, josta saan paljon verkkokauppaan liittyvää dataa minkä analysointi on elintärkeää menestymiseni

kannalta. Nykyiset verkkokauppaohjelmat ja googlen työkalut visualisoivat datan valmiiksi, mutta kurssin opeilla saan kerättyä tiedot yhteen helposti luettavaksi kokonaisuudeksi.

Suurena apuna oppimisprosessissa oli erittäin hyvät ryhmätyökaverit, joilta sai apua ja tukea, kun oma ymmärrys loppui. Jyrki on analyttisin koskaan tapaamani kaveri ja lisäksi todella taitava Excelin käyttäjä. Janne taas ymmärtää ja hallitsee kokonaisuuksia siten, että ryhmätyön tekeminen oli helppoa alusta asti.

Aineistot

Säähavainnot, Ilmatieteenlaitos. Osoite: <https://www.ilmatieteenlaitos.fi/havaintojen-lataus#!/>

Korkeasaaren kävijämäärät, Pääkaupunkiseudun avoimen datan palvelu. Osoite: <https://hri.fi/data/dataset//korkeasaaren-kavijamaarat>

Korkeasaaren aukioloajat. Sähköposti, katso liite 3 ja liite 4.

Lähteet

Bojic, P., Greasley, A. ja Hickie, S. 2015. Business Information Systems. Technology, Development and Management for the E-Business. Fifth edition.

Rasch, R. ja Schott, D. 2018. Mathematical Statistics. Wiley. E-kirja. Luettavissa: <https://haaga-helia.finna.fi/>. Luettu: 19.2.2021.

Stolwijk, A. M., Straatman, H. ja Zielhuis, G. A., 1999. Studying seasonality by using sine and cosine functions in regression analysis. Journal of Epidemiology & Community Health, numero 53(4), sivut 235-238. Luettavissa: <https://jech.bmj.com/content/53/4/235.short>. Luettu: 19.2.2021.

Liitteet

Liite 1: Korkeasaaren kävijämäärät

Lähde: <https://hri.fi/data/fi/dataset/korkeasaaren-kavijamaarat/resource/1690b031-0a55-4d12-852e-ba1bc36ead6f>

Excel tiedostot: Kävijämäärä 2015–2019

Tiedostojen Kävijämäärä 2015–2016 tiedostojen rakenne oli seuraavanlainen

Osa 1 / 2:

KORKEASAAREN ELÄINTARHA										KÄVIJÄTILASTO 2016										HEINÄKUU									
MUSTIKKAMAA										KAUPPATORI																			
MAKSETUT KÄYNNIT					ILMAISKÄYNNIT					PERHELIPPU					MAKSETUT KÄYNNIT					ILMAISKÄYNNIT					PERHELIPPU				
Aik.	Lapset	HK	VK		Aik.	Lapset				Aik.	Lapset				Aik.	Lapset	HK	VK		Aik.	Lapset				Aik.	Lapset			
1 pe	1180	545	3	119	63	313		134	183	2540	471	244	3	12	9	105	38	58		940									
2 la	2311	713	0	151	41	402		164	227	4009	874	365	4	20	6	129	64	95		1557									
3 su	651	242	0	116	37	110		28	76	1260	218	73	7	5	0	31	11	15		360									
4 ma	786	346	7	122	19	166		72	92	1610	208	104	4	5	0	35	25	37		418									
5 ti	1283	650	0	157	36	271		121	173	2691	473	234	0	10	1	99	45	65		927									
6 ke	2369	1082	5	286	30	547		287	396	5001	807	431	0	5	4	149	65	93		1554									
7 to	355	149	0	92	36	73		44	62	811	109	66	0	0	0	11	2	3		191									
8 pe	1469	608	5	111	27	297		157	229	2902	493	244	3	14	1	88	33	50		926									
9 la	2942	1015	0	152	194	725		138	207	5373	942	358	4	8	7	162	49	72		1602									
10 su	1359	454	2	111	357	85		66	99	2533	357	146	1	4	11	66	31	46		662									
11 ma	1605	752	1	71	49	376		185	259	3298	616	352	6	4	1	110	35	48		1172									
12 ti	1829	773	12	138	73	364		201	291	3671	701	390	8	7	5	152	46	67		1376									
13 ke	2436	1029	5	185	58	450		284	402	4849	987	455	1	3	3	208	68	102		1827									
14 to	2466	1029	3	160	67	472		302	440	4939	853	409	2	15	14	154	73	109		1629									
15 pe	214	116	0	14	6	46		44	63	503	49	33	2	4	1	2	6	8		105									
16 la	2973	976	1	154	36	529		282	402	5353	1003	350	4	4	10	176	76	112		1735									
17 su	1830	622	2	151	22	287		217	314	3445	489	181	8	9	3	76	37	55		858									
18 ma	2276	907	4	118	26	410		240	341	4322	786	378	2	10	6	154	49	75		1460									
19 ti	2426	954	5	164	31	427		275	387	4669	949	468	0	18	24	194	84	129		1866									
20 ke	2356	966	4	110	80	408		257	362	4543	1010	465	9	11	16	208	97	145		1961									
21 to	2085	843	4	104	48	440		239	333	4096	736	365	5	7	6	157	50	82		1408									

Osa 2 / 2:

KÄYNNIT YHTEENSÄ																			
HAKANIEMI										YHTEENSÄ									
MAKSETUT KÄYNNIT					ILMAISKÄYNNIT					PERHELIPPU					YHTEENSÄ				
Aik.	Lapset	HK	VK		Aik.	Lapset				Aik.	Lapset								
0	0	0	0		0	0		0	0	0	0		0	0	3480				
0	0	0	0		0	0		0	0	0	0		0	0	5566				
0	0	0	0		0	0		0	0	0	0		0	0	1620				
60	39	1	0		0	6		8	11	125					2153				
0	0	0	0		0	0		0	0	0	0		0	0	3618				
0	0	0	0		0	0		0	0	0	0		0	0	6555				
30	22	2	0		0	4		4	6	68					1070				
0	0	0	0		0	0		0	0	0	0		0	0	3828				
0	0	0	0		0	0		0	0	0	0		0	0	6975				
0	0	0	0		0	0		0	0	0	0		0	0	3195				
0	0	0	0		0	0		0	0	0	0		0	0	4470				
0	0	0	0		0	0		0	0	0	0		0	0	5047				
0	0	0	0		0	0		0	0	0	0		0	0	6676				
0	0	0	0		0	0		0	0	0	0		0	0	6568				
0	0	0	0		0	0		0	0	0	0		0	0	608				
0	0	0	0		0	0		0	0	0	0		0	0	7088				
0	0	0	0		0	0		0	0	0	0		0	0	4303				
0	0	0	0		0	0		0	0	0	0		0	0	5782				
0	0	0	0		0	0		0	0	0	0		0	0	6535				
0	0	0	0		0	0		0	0	0	0		0	0	6504				
0	0	0	0		0	0		0	0	0	0		0	0	5504				
0	0	0	0		0	0		0	0	0	0		0	0	4863				
0	0	0	0		0	0		0	0	0	0		0	0	5568				
0	0	0	0		0	0		0	0	0	0		0	0	4652				
0	0	0	0		0	0		0	0	0	0		0	0	4297				
0	0	0	0		0	0		0	0	0	0		0	0	4991				
0	0	0	0		0	0		0	0	0	0		0	0	4917				

Tiedostojen Kävijämäärä 2017–2019 tiedostojen rakenne oli puolestaan seuraavanlainen

KORKEASAAREN ELÄINTARHA		KÄVIJÄTILASTO 2017
Tammikuu		Käynnit yhteensä
1 su		538
2 ma		510
3 ti		322
4 ke		204
5 to		70
6 pe		103
7 la		86
8 su		299
9 ma		66
10 ti		58
11 ke		25
12 to		19
13 pe		48
14 la		259
15 su		336
16 ma		38
17 ti		105
18 ke		99
19 to		39
20 pe		146

Liite 2: Ilmatieteenlaitoksen säähavainnot

Lähde: <https://www.ilmatieteenlaitos.fi/havaintojen-lataus#!/>

Havaintoasemat ja ajat:

Helsinki Kumpula, Asemakoodi 101004, aikaväli 1.1.2015 – 31.12.2019

Helsinki Kaisaniemi, Asemakoodi 100971, aikaväli 2.11.2017 – 6.6.2018

Excel tiedostojen rakenne

Osa 1 / 2:

sep=	Vuosi	Kk	Pv	Klo	Aikavyöhyke	Pilvien määrä (1/8)	Ilmanpaine (msl) (hPa)	Sademäärä (mm)	Suhteellinen kosteus (%)	Sateen intensiteetti (mm/h)	Lumensyvyys (cm)
2015	1	1	00:00:00	UTC		7	1010,7	0	89	0	10
2015	1	1	00:57:36	UTC		7	1010,5	0	89	0	9
2015	1	1	01:55:12	UTC		5	1010,5	0	90	0	10
2015	1	1	03:07:12	UTC		7	1010,1	0	91	0	9
2015	1	1	04:04:48	UTC		7	1009,7	0	89	0	9
2015	1	1	05:02:24	UTC		7	1009,4	0	89	0	10
2015	1	1	06:00:00	UTC		7	1009,2	0	86	0	9
2015	1	1	06:57:36	UTC		7	1008,8	0	90	0	9
2015	1	1	07:55:12	UTC		7	1008,4	0	91	0	9
2015	1	1	09:07:12	UTC		8	1008,1	0	94	0	9
2015	1	1	10:04:48	UTC		9	1006,9	0	95	0	8
2015	1	1	11:02:24	UTC		8	1006,6	0	96	0	8
2015	1	1	12:00:00	UTC		7	1005,9	0	96	0	8
2015	1	1	12:57:36	UTC		8	1005,3	0	96	0	8
2015	1	1	13:55:12	UTC		8	1004,6	0	96	0	8
2015	1	1	15:07:12	UTC		8	1004,4	0	92	0	7
2015	1	1	16:04:48	UTC		8	1003,7	0	90	0	7
2015	1	1	17:02:24	UTC		8	1003	0	90	0	7
2015	1	1	18:00:00	UTC		8	1002,9	0	89	0	7
2015	1	1	18:57:36	UTC		8	1002,4	0	90	0	6
2015	1	1	19:55:12	UTC		8	1001,9	0	89	0	6
2015	1	1	21:07:12	UTC		8	1001,6	0	87	0	6
2015	1	1	22:04:48	UTC		8	1001	0	87	0	5
2015	1	1	23:02:24	UTC		8	1000,2	0	87	0	5
2015	1	2	00:00:00	UTC		8	999,4	0	90	0	5
2015	1	2	00:57:36	UTC		8	998,3	0	91	0	5

Osa 2 / 2:

Ilman lämpötila (degC)	Kastepistelämpötila (degC)	Näkyvyys (m)	Tuulen suunta (deg)	Puuskanopeus (m/s)	Tuulen nopeus (m/s)
3,3	1,6	13700	259	7,1	4,4
3,2	1,5	17390	254	6,5	4,6
3	1,6	18170	251	6,3	4,3
2,6	1,3	18620	240	6,4	4,2
3,3	1,7	30320	236	5,4	3,7
3,4	1,8	36370	240	7,1	5,2
3,7	1,5	46960	251	5,9	3,9
3,1	1,6	33600	209	2,7	2,2
2,7	1,4	35590	208	5,5	3,7
2,4	1,5	1790	215	7,7	5,3
2,3	1,6	1050	219	5,6	3,5
1,8	1,2	3970	227	9,3	6,2
1,9	1,3	6220	237	7,9	5,5
1,9	1,3	8660	226	7,1	4,7
1,9	1,3	9940	229	8,4	5,1
2,2	1	13310	234	8,9	5,6
2,4	0,9	14570	238	7,4	5,2
2,8	1,3	16640	238	7,4	4,8
3,2	1,5	18870	240	9,4	6,9
3,4	2	12570	244	9,8	6,2
3,7	2	33570	246	7,6	4,8
4	1,9	49310	242	8,3	5,8
3,9	1,9	49490	242	8,3	6,1
3,9	2	49470	232	6,6	4,5
3,5	2	37770	221	6,4	4,3
3,6	2,2	29510	226	7	4,5
3,8	2,5	26990	229	6,2	4,1

Liite 3: Korkeasaaren sähköposti

Huomio! Viestiketjusta poistettu kaikki sähköpostiosoitteet, puhelinnumerot ja linkit

Korkeasaari

ti 26.1.2021 11.21

– **Vastaanottaja:** Aho Jyrki
AUKIOLO 2015-19 SÄHKÖPOSTIKYSELY.docx

14 kt

Hei Jyrki!

Mielenkiintoinen tutkimus teillä meneillään 😊

Ainakin näin parin vuoden työkokemuksen kerryttämällä näppituntumalla voisin sanoa, että sää vaikuttaa kävijätilastoihin hyvinkin paljon.

Laitoin liitteeksi tiedoston, johon haalin kokoon kaiken tiedon mitä sain vanhoista työvuorolis-toista yms. koottua, toivottavasti siitä on teille apua.

Tsemppiä tutkimuksen tekoon! 😊

Ystävällisin terveisin,

Viliina



Korkeasaaren lipunmyynti
Helsinki Zoo ticket sales



Lähetäjä: Aho Jyrki

Lähetetty: maanantai 25. tammikuuta 2021 12:04

Vastaanottaja: ~~Korkeasaari~~

Aihe: Tietoa tutkimusta kouluprojektia varten

Hei

Olemme Haaga-Helian opiskelijoita ja ajattelimme suorittaa tutkimuksen sään vaikutuksesta Korkeasaaren kävijämääriin. Meille muodostui ongelmaksi, ettei meillä ole tietoa teidän aukioloista (ja poikkeusaukioloajoista) vuosien 2015 - 2019 vuosien ajalta. Olisiko tätä kautta mahdollista saada nämä tiedot?

Sääolosuhteet ja Korkeasaaren kävijämäärät meillä on jo tiedossa.

Ystävällisin terveisin

Jyrki Aho

Hanki Outlook for Android

Liite 4: Korkeasaaren sähköpostin liitetiedosto

2015

1.1.-31.3. 10-16
1.-30.4. 10-18
1.5.-31.8. 10-20
1.-30.9. 10-18
1.10.-31.12. 10-16

poikkeukset 2015

8.-9.2., 21.-22.2. 10-18 talvireitti
26.4. 9-18 partiolaiset saarella
19.6. 10-17 juhannusaatto
4. & 11.9. 16-00 Kissojen yöt
4.10. 10-18 ilmapäivä
1.11. 9-16 partiotapahtuma
24.12. suljettu

2016

1.1.-31.3. 10-16
1.-30.4. 10-18
1.5.-31.8. 10-20
1.-30.9. 10-18
1.10.-31.12. 10-16

poikkeukset 2016

10.-13., 16.-20. & 23.-27.5 opastettuja
kouluryhmiä jo ennen klo 9:30
24.6. 10-17 juhannusaatto
2. & 9.9. 16-00 Kissojen yöt
24.12. suljettu

2017

1.1.-31.3. 10-16
1.-30.4. 10-18
1.5.-31.8. 10-20
1.-30.9. 10-18
1.10.-31.12. 10-16

poikkeukset 2017

17.6. 10-22 Suomi 100 kesäyö
23.6. 10-17 juhannusaatto
1. & 8.9. 16-00 Kissojen yöt
4.10. 10-17 ilmapäivä
6.9. 16-22 Viattelysten ilta
24.12 suljettu

2018

1.1.-31.3. 10-16
1.-30.4. 10-18
1.5.-31.8. 10-20
1.-30.9. 10-18
1.10.-31.12. 10-16

poikkeukset 2018

30.3-2.4. 10-18 pääsiäinen
22.6. 10-17 juhannusaatto
7.9. & 14.9. 16-00 Kissojen yöt
12.9. 10-22 Viettelysten ilta
24.12. suljettu

2019

1.1.-28.2. 10-16
1.3.-30.4. ma-pe 10-16, la-su 10-18
1.-31.5. ma-pe 10-18, la-su 10-20
1.6.-31.8. 10-20
1.-30.9. 10-18
1.-31.10. ma-pe 10-16, la-su 10-18
1.11.-31.12. 10-16

poikkeukset 2019

16.-24.2. 10-18 talviloma
19.-22.4. 10-18 pääsiäinen
1.5. & 30.5. 10-20 vappu & helatorstai
21.6. 10-17 juhannusaatto
6. & 13.9. 16-00 Kissojen yöt
11.9. 10-22 Viettelysten ilta
4.10. 10-18 ilmapäivä
14.-18.10. 10-18 syysloma
26.-27.10. 10-18 halloween
2.-3.11. 10-18 halloween
24.12. 10-14
25.-31.12. 10-18 joululoma & LUX

Liite 5: Kooste Regressiossa käytetystä datasta

Alla esimerkki koostetusta datasta, missä säähavainnot vastaavat yhden päivän keskiarvoa.

Osa 1 / 3:

Pvm	V_paiva	Viikonpaiva	Pvm2	Pvm3	sin_1pt	cos_1pt	sin_2pt	cos_2pt
2015-01-01	5	Torstai	0	0	0	0	1	0
2015-01-02	6	Perjantai	0.002739726027397	0.002739726027397	0.008606996888688	0.999962959116266	0.017213356155835	0.999851839209116
2015-01-03	7	Lauantai	0.005479452054795	0.005479452054795	0.017213356155835	0.999851839209116	0.034421611622746	0.999407400739705
2015-01-04	1	Sunnuntai	0.008219178082192	0.008219178082192	0.025818440227133	0.999666648510511	0.051619667223254	0.998666816288476
2015-01-05	2	Maanantai	0.010958904109589	0.010958904109589	0.034421611622746	0.999407400739705	0.06880242680232	0.997630305306586
2015-01-06	3	Tiistai	0.013698630136986	0.013698630136986	0.043022233004531	0.99907411510223	0.085964798737447	0.996298174934608
2015-01-07	4	Keskiviikko	0.016438356164384	0.016438356164384	0.051619667223254	0.998666816288476	0.103101697447435	0.994670819911521
2015-01-08	5	Torstai	0.019178082191781	0.019178082191781	0.060213277365793	0.998185534471859	0.120208044899353	0.992748722457774
2015-01-09	6	Perjantai	0.021917808219178	0.021917808219178	0.06880242680232	0.997630305306586	0.137278772113265	0.990532452132223
2015-01-10	7	Lauantai	0.024657534246575	0.024657534246575	0.077386479233463	0.997001169925015	0.154308820664281	0.988022665663697
2015-01-11	1	Sunnuntai	0.027397260273973	0.027397260273973	0.085964798737447	0.996298174934608	0.171293144181478	0.985220106756061
2015-01-12	2	Maanantai	0.03013698630137	0.03013698630137	0.094536749817199	0.995521372414475	0.188226709843244	0.982125605868001
2015-01-13	3	Tiistai	0.032876712328767	0.032876712328767	0.103101697447435	0.994670819911521	0.205104499866619	0.978740079966915
2015-01-14	4	Keskiviikko	0.035616438356164	0.035616438356164	0.111659007121694	0.993746580436178	0.221921513004165	0.957064532257195
2015-01-15	5	Torstai	0.038356164383562	0.038356164383562	0.120208044899353	0.99274872245774	0.23867276600595	0.97110005188295
2015-01-16	6	Perjantai	0.041095890410959	0.041095890410959	0.128748177452581	0.99167731989929	0.255353295116187	0.966847813605278
2015-01-17	7	Lauantai	0.043835616438356	0.043835616438356	0.137278772113265	0.990532452132223	0.271958157534105	0.962309077454149
2015-01-18	1	Sunnuntai	0.046575342465754	0.046575342465754	0.145799196919875	0.989314203970366	0.288482432880609	0.957485188355039
2015-01-19	2	Maanantai	0.049315068493151	0.049315068493151	0.154308820664281	0.988022665663697	0.304921224656289	0.952377575730398
2015-01-20	3	Tiistai	0.052054794520548	0.052054794520548	0.162807012938517	0.986657932891657	0.321269661692364	0.946987753076075
2015-01-21	4	Keskiviikko	0.054794520547945	0.054794520547945	0.171293144181478	0.985220106756061	0.337522899594113	0.941317317512847
2015-01-22	5	Torstai	0.057534246575343	0.057534246575343	0.179766585725562	0.98370929377361	0.353676122176372	0.935367949313148

Osa 2 / 3:

sin_3pt	cos_3pt	sin4_pt	cos_4pt	poutaisia_pai	epavakaisia_pai	sateisia_pai	Pilvien_maara	sademaara
0	0	1	0	0	0	0	0	8 0.00
0.025818440227133	0.999666648510511	0.034421611622746	0.999407400739705	0	1	0	0	8 0.75
0.051619667223254	0.998666816288476	0.06880242680232	0.997630305306586	1	0	0	0	8 0.15
0.077386479233463	0.997001169925015	0.103101697447435	0.994670819911521	2	0	0	0	7 0.00
0.103101697447435	0.994670819911521	0.137278772113265	0.990532452132223	3	0	0	0	0 0.00
0.128748177452581	0.99167731989929	0.171293144181478	0.985220106756061	4	0	0	0	0 0.00
0.154308820664281	0.988022665663697	0.205104499866619	0.978740079966915	5	0	0	0	6 0.00
0.179766585725562	0.98370929377361	0.23867276600595	0.97110005188295	0	1	0	0	8 0.67
0.205104499866619	0.978740079966915	0.271958157534105	0.962309077454149	1	0	0	0	7 0.00
0.230305670230612	0.973118337233262	0.304921224656289	0.952377575730398	2	0	0	0	8 0.00
0.255353295116187	0.966847813605278	0.337522899594113	0.941317317512847	3	0	0	0	8 0.02
0.280230675199216	0.959932689659745	0.369724542890673	0.929141411403174	4	0	0	0	4 0.00
0.304921224656289	0.952377575730398	0.401487989205973	0.915864288267287	5	0	0	0	7 0.00
0.32940848222453	0.944187508834199	0.432775592550431	0.901501684131884	6	0	0	0	7 0.00
0.353676122176372	0.935367949313148	0.463550270902851	0.886070621534138	7	0	0	0	4 0.00
0.37770965203965	0.92592477719385	0.493775550159977	0.869589389346611	8	0	0	0	8 0.07
0.401487989205973	0.915864288267287	0.52341560736555	0.852077521101309	9	0	0	0	8 0.02
0.425000339969554	0.905193189891397	0.55243531316762	0.833555717183857	10	0	0	0	8 0.00
0.448229341740411	0.893918596519257	0.580800273453801	0.814046093508218	11	0	0	0	7 0.00
0.471159507673864	0.882048024955854	0.608476870115126	0.793571608952147	12	0	0	0	8 0.00
0.493775550159977	0.869589389346611	0.635432300890177	0.772156584499164	13	0	0	0	8 0.00
0.516062391015853	0.856550995901003	0.661634618242278	0.749826401204569	14	0	0	0	8 0.00

Osa 3 / 3:

lampotila	Lampotila^2	exp(lampotila/10)	tuulennopeus	suhteellinenkestausvieraita	Vieraita_per_h	ln(vieraita_per_h)	vierasluokka
2.0000	4.0000	1.2214	5.1	95.1666666666667	381 63.5	1.80277372529198	ErittainPaljon
4.3167	18.6336	1.5398	9.9	93.1666666666667	86 14.3333333333333	1.15634720085992	Normaali
1.3667	1.8678	1.1464	8.5	89.1666666666667	250 41.6666666666667	1.61978875828839	Paljon
-1.3333	1.7778	0.8752	6.0	78.5	433 72.1666666666667	1.85833664596972	ErittainPaljon
-10.0667	101.3378	0.3654	4.5	75.8333333333333	327 54.5	1.73639650227664	ErittainPaljon
-11.1333	123.9511	0.3285	0.7	73.8333333333333	293 48.8333333333333	1.68871636997047	ErittainPaljon
-2.2333	4.9878	0.7998	8.4	71.1666666666667	134 22.3333333333333	1.34895354798116	Normaali
1.3167	1.7336	1.1407	7.0	94.6666666666667	137 22.8333333333333	1.35856931677276	Normaali
0.0833	0.0069	1.0084	2.8	96	127 21.1666666666667	1.32565247057231	Normaali
-1.5000	2.2500	0.8607	2.2	77.1666666666667	289 48.1666666666667	1.6827465923729	ErittainPaljon
-8.0500	64.8025	0.4471	7.7	87.5	64 10.6666666666667	1.02802872360024	Vahainen
-10.9667	120.2678	0.3340	2.3	81	25 4.16666666666667	0.619788758288394	Vahainen
1.2833	1.6469	1.1369	7.4	73.8333333333333	16 2.66666666666667	0.425968732272281	Vahainen
2.4667	6.0844	1.2798	7.4	88.1666666666667	36	6 0.778151250383644	Vahainen
2.4167	5.8403	1.2734	6.0	79.6666666666667	41 6.83333333333333	0.834632606336092	Vahainen
3.1333	9.8178	1.3680	9.9	88.6666666666667	13 2.16666666666667	0.335792101923193	Vahainen
2.5833	6.6736	1.2948	6.1	90	182 30.3333333333333	1.48192013760143	Normaali
2.0833	4.3403	1.2316	5.2	82.1666666666667	213 35.5	1.55022835305509	Paljon
0.0167	0.0003	1.0017	2.7	87.8333333333333	34 5.66666666666667	0.753327666658611	Vahainen
-2.2667	5.1378	0.7972	4.3	77.5	19 3.16666666666667	0.500602350569185	ErittainVahainen
-2.1000	4.4100	0.8106	7.1	77.8333333333333	32 5.33333333333333	0.726998727936262	ErittainVahainen

Liite 6: Päättäntäpuun muodostamisessa käytetty data

Alla arff-muotoisen tiedosto ja sen kaksikymmentä ensimmäistä riviä.

@relation korkeasaari

@attribute paivamaara date yyyy-mm-dd

@attribute viikonpaiva {Maanantai, Tiistai, Keskiviikko, Torstai, Perjantai, Lauantai, Sunnuntai}

@attribute sademaara numeric

@attribute lampotila numeric

@attribute tuulennopeus numeric

@attribute vieraitaavgh numeric

@attribute vieraitamaara numeric

@attribute vieraskorjattu numeric

@attribute vieraita {Paljon, Normaali, Vahainen, ErittainPaljon, ErittainVahainen}

@data

2015-01-01,Torstai,0.00,2.0,5.1,2.82,381,319.5723986,ErittainPaljon

2015-01-02,Perjantai,0.75,4.3,9.9,2.85,86,24.36539357,Normaali

2015-01-03,Lauantai,0.15,1.4,8.5,2.88,250,187.7639077,Paljon

2015-01-04,Sunnuntai,0.00,-1.3,6.0,2.91,433,369.7686866,ErittainPaljon

2015-01-05,Maanantai,0.00,-10.1,4.5,2.94,327,262.3809622,ErittainPaljon

2015-01-06,Tiistai,0.00,-11.1,0.7,2.96,293,226.6024501,ErittainPaljon

2015-01-07,Keskiviikko,0.00,-2.2,8.4,2.99,134,65.43534415,Normaali

2015-01-08,Torstai,0.67,1.3,7.0,3.03,137,65.88231057,Normaali

2015-01-09,Perjantai,0.00,0.1,2.8,3.06,127,52.94648063,Normaali

2015-01-10,Lauantai,0.00,-1.5,2.2,3.09,289,211.6314419,ErittainPaljon

2015-01-11,Sunnuntai,0.02,-8.1,7.7,3.12,64,-17.05877158,Vahainen

2015-01-12,Maanantai,0.00,-11.0,2.3,3.15,25,-60.11969017,Vahainen

2015-01-13,Tiistai,0.00,1.3,7.4,3.18,16,-73.54642087,Vahainen

2015-01-14,Keskiviikko,0.00,2.5,7.4,3.22,36,-58.33366041,Vahainen

2015-01-15,Torstai,0.00,2.4,6.0,3.25,41,-58.47570935,Vahainen

2015-01-16,Perjantai,0.07,3.1,9.9,3.28,13,-91.96648713,Vahainen

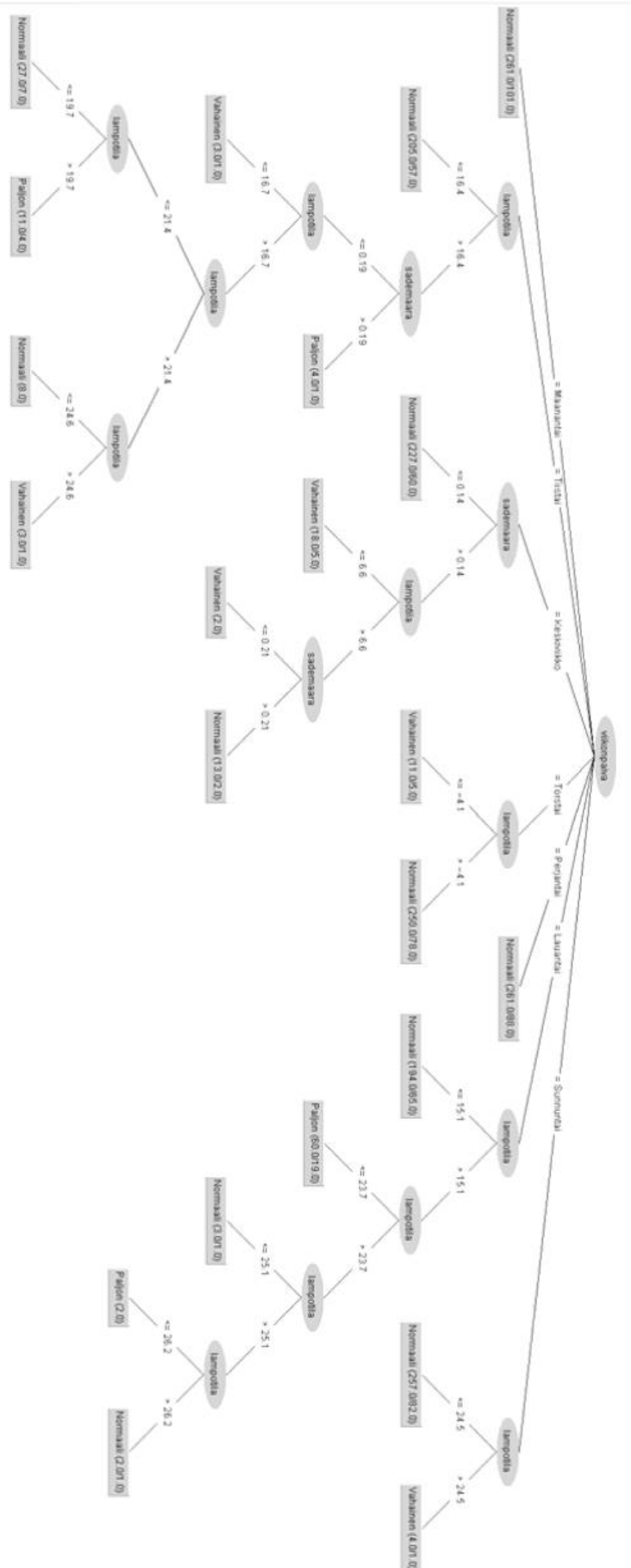
2015-01-17,Lauantai,0.02,2.6,6.1,3.32,182,71.20045206,Normaali

2015-01-18,Sunnuntai,0.00,2.1,5.2,3.35,213,96.03190264,Paljon

2015-01-19,Maanantai,0.00,0.0,2.7,3.39,34,-89.46500971,Vahainen

2015-01-20,Tiistai,0.00,-2.3,4.3,3.42,19,-111.282846,ErittainVahainen

Liite 7: Päättöspuu



Liite 8: Regressiomalli ja logistinen malli

Apufunktiot

Malleihin liittyvät apufunktiot

$$T = (\text{päivien lkm vuoden alusta})/365 + (\text{vuosiluku} - 2015)$$

$$I_A(z) = \begin{cases} 1, & \text{kun } z \in A \\ 0, & \text{kun } z \notin A \end{cases}$$

Sinimalli

$$s(T) = 0,59783136 * \sin(-6,30897004 * T - 1,50974012) + 1,87774338$$

Ajalla varustettu regressiomalli

vierailijoita per h =

$$\begin{aligned} & 24,197719720979 * (1 - I_{\{Ma\}}(X_{\text{viikonpaiva}})) + 92,845459864097 * I_{\{La\}}(X_{\text{viikonpaiva}}) \\ & + 48,314935807555 * I_{\{Su\}}(X_{\text{viikonpaiva}}) \\ & + 6,987346878689 * \sin(\pi * T) - 13,673580644725 * \sin(2 * \pi * T) \\ & - 113,770946721921 * \cos(2 * \pi * T) - 9,519554518738 * \sin(3 * \pi * T) \\ & + 19,170031383178 * \sin(4 * \pi * T) + 62,839500829712 * \cos(4 * \pi * T) \\ & - 55,129929919174 * X_{\text{sademaara}} - 4,580082905769 * X_{\text{tuulennopeus}} \\ & - 1,379035397032 * X_{\text{suhteellinenkosteus}} \\ & + 3,337208763289 * X_{\text{lampotila}} - 4,887411857058 * \exp(X_{\text{lampotila}}/10) \\ & + 217,79749512408 \end{aligned}$$

Ilman aikaa varustettu regressiomalli

vieraita per h =

$$\begin{aligned} & 22,8333 * (1 - I_{\{Ma\}}(X_{\text{viikonpaiva}})) + 89,6179 * I_{\{La\}}(X_{\text{viikonpaiva}}) \\ & + 48,8995 * I_{\{Su\}}(X_{\text{viikonpaiva}}) - 29,2368 * X_{\text{sademara}} \\ & + 6,2892 * X_{\text{lampotila}} + 13,3617 * \exp(X_{\text{lampotila}} / 10) \\ & - 5,7167 * X_{\text{tuulennopeus}} + 1,6598 * X_{\text{suhteellinenkosteus}} \\ & + 158,7453 \end{aligned}$$

Liite 9: Logistisen mallit

=== Classifier model (full training set) ===

Logistic Regression with ridge parameter of 1.0E-8

Coefficients...

Variable	Class			
	Paljon	Normaali	Vahainen	ErittainPaljon
=====				
viikonpaiva=Maanantai	-1.6916	-1.1964	-0.4555	-1.7591
viikonpaiva=Tiistai	-0.5503	-0.7215	-0.1944	-0.833
viikonpaiva=Keskiviikko	-0.5668	-0.374	-0.1395	-0.8741
viikonpaiva=Torstai	-0.704	-0.3184	0.0994	-0.6271
viikonpaiva=Perjantai	-0.7395	-0.3061	-0.1141	-0.6847
viikonpaiva=Lauantai	2.5447	1.898	0.5831	3.2672
viikonpaiva=Sunnuntai	1.7055	1.0172	0.2205	1.508
pvm3	0.0515	0.1143	0.0561	0.08
sin_1pt	0.1194	0.0884	0.1032	0.1247
cos_1pt	-0.0706	0.0032	-0.0479	0.0661
sin_2pt	-0.3934	-0.3272	-0.2917	-0.0569
cos_2pt	1.7399	1.8845	1.8062	2.4128
sin_3pt	-0.0367	-0.0401	0.074	-0.1261
cos_3pt	0.1372	-0.0883	-0.0044	0.0115
sin_4pt	0.02	-0.2952	-0.1307	0.3326
cos_4pt	0.4082	-0.4826	-0.7629	0.5909
sademaara	-1.8727	-1.6761	-0.4131	-4.2964
lampotila	0.191	0.104	0.0962	0.2697
lampotila_exp	-0.4148	-0.1852	-0.1409	-0.5595
tuulennopeus	-0.1568	-0.1349	-0.0485	-0.1904
suhteellinenkosteus	-0.0716	-0.0449	-0.0409	-0.0755
Intercept	6.235	4.7689	4.5183	6.3933

Odds Ratios...

Variable	Class			
	Paljon	Normaali	Vahainen	ErittainPaljon
=====				
viikonpaiva=Maanantai	0.1842	0.3023	0.6341	0.1722
viikonpaiva=Tiistai	0.5768	0.486	0.8233	0.4347
viikonpaiva=Keskiviikko	0.5674	0.688	0.8698	0.4172
viikonpaiva=Torstai	0.4946	0.7273	1.1046	0.5341
viikonpaiva=Perjantai	0.4774	0.7363	0.8922	0.5042
viikonpaiva=Lauantai	12.7396	6.6728	1.7916	26.2385
viikonpaiva=Sunnuntai	5.5044	2.7654	1.2467	4.5178
pvm3	1.0529	1.1211	1.0577	1.0833
sin_1pt	1.1268	1.0924	1.1087	1.1328
cos_1pt	0.9318	1.0032	0.9533	1.0683
sin_2pt	0.6747	0.7209	0.747	0.9447
cos_2pt	5.697	6.5834	6.0874	11.1657
sin_3pt	0.964	0.9607	1.0768	0.8815
cos_3pt	1.1471	0.9155	0.9956	1.0116

sin_4pt	1.0202	0.7444	0.8775	1.3946
cos_4pt	1.5041	0.6172	0.4663	1.8057
sademaara	0.1537	0.1871	0.6616	0.0136
lampotila	1.2105	1.1096	1.101	1.3096
lampotila_exp	0.6605	0.8309	0.8686	0.5715
tuulennopeus	0.8549	0.8738	0.9527	0.8266
suhteellinenkosteus	0.9309	0.9561	0.9599	0.9273

Time taken to build model: 0.4 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	874	47.8642 %
Incorrectly Classified Instances	952	52.1358 %
Kappa statistic	0.24	
Mean absolute error	0.2621	
Root mean squared error	0.3643	
Relative absolute error	88.1132 %	
Root relative squared error	94.4684 %	
Total Number of Instances	1826	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0,156	0,057	0,295	0,156	0,204	0,131	0,723	0,271	Paljon
0,176	0,090	0,357	0,176	0,235	0,114	0,604	0,296	Normaali
0,850	0,515	0,522	0,850	0,647	0,344	0,725	0,591	Vahainen
0,381	0,069	0,478	0,381	0,424	0,344	0,769	0,410	ErittainPaljon
0,250	0,036	0,449	0,250	0,321	0,279	0,785	0,364	ErittainVahainen
Weighted Avg.								
0,479	0,246	0,441	0,479	0,431	0,258	0,711	0,434	

=== Confusion Matrix ===

a	b	c	d	e	<-- classified as
38	42	96	54	13	a = Paljon
31	71	252	37	13	b = Normaali
20	45	618	17	27	c = Vahainen
31	36	88	99	6	d = ErittainPaljon
9	5	130	0	48	e = ErittainVahainen