

Hanna Kaimo – Saku Ihalainen – Elyas Addawe
Projekti (CRISP-DM)
Helsingin kaupunkipyöräjärjestelmä

Metropolia Ammattikorkeakoulu
Tieto- ja viestintätekniikka
Datan käsittely ja koneoppiminen
22.4.2022

1 Tavoitteet

Helsingin ja Espoon kaupunkipyöräjärjestelmä avattiin toukokuussa v. 2016. Silloin Helsingissä oli 50 kaupunkipyöräasemaa. Viime vuonna niitä oli 352.

Helsingissä ajettiin huhtikuussa 2021 keskimäärin 8 208 kaupunkipyörämatkaa päivässä. Vertailuna esimerkiksi New York:n kaupunkipyöräajomäärä oli tämän vuoden helmikuussa 42 682 ajettua matkaa päivässä. Joten suhteutettuna asukasmääriin (Helsinki n. 631 200 asukasta, New York n. 8 419 000 asukasta) Helsingissä ajetaan todella paljon kaupunkipyöillä, joten niiden järjestelmän kehittäminen ja ylläpitäminen tutkimustulosten pohjalta kannattaa ainakin tällä hetkellä.

Kaupunkipyörällä saa ajaa puoli tuntia perusmaksulla (päivä, viikko tai koko kausi(35 €)), jonka jälkeen veloitetaan lisämaksu. Tällä pyritään siihen, että käyttöajat olisivat lyhyitä ja pyöriä riittäisi kaikkien käyttöön.

Tavoitteenamme on selvittää kuinka käytettyjä eri pyöräasemat ovat: kannattaako pyöriä niihin laittaa lisää tai jos asemaa on käytetty todella vähän, voidaanko se poistaa? Ovatko pyörien käyttöajat tarpeeksi matalia? Ja ylipäättäään onko järjestelmä toimiva ja mitä muutoksia datan perusteella kaupunkipyöräjärjestelmään pitäisi tehdä, jotta se olisi vielä kannattavampaa ja päivitetty nykyaikaan/ ajan tasalle. Ja toimiiko järjestelmä oikein?

2 Data

Data on otettu HSL:n avoin data- sivulta kaupunkipyöillä ajettujen matkojen kohdalta. Daten omistaa City Bike Finland.

Attribuutit:

Departure (date and time): nominaali- muuttuja, vaihteluväli 1.4 klo 5:06:37 - 30.4.2021 klo 23:59:54

Return (date and time): nominaali- muuttuja, vaihteluväli 1.4 klo 6:05:39 - 1.5.2021 klo 8:15:54

DepartureStationID: numeerinen muuttuja, vaihteluväli 1 - 997

DepartureStationName: nominaali- muuttuja, vaihteluväli 1 - 352

ReturnStationID: numeerinen muuttuja, vaihteluväli 1 - 997

ReturnStationName: nominaali- muuttuja, vaihteluväli 1 - 352

CoveredDistance (m): numeerinen muuttuja, vaihteluväli (-9) - 153160

Duration (sec): numeerinen muuttuja, vaihteluväli 0 - 5969

Datasta saattavat puuttua pyörät, joita on käytetty virheellisesti, esim. matkan päättyessä pyörän palautus ei ole rekisteröitynyt järjestelmään, koska se on tehty virheellisesti. Negatiivinen arvo pyöräilymatkassa kannattaa poistaa. Myös pisimmät matkat ovat hieman epäilyttäviä esim. 153,160 km.

Datan koko on 246 226 arvoa. Se on kattava.

3 Daten valmistelu

Datassa oli turhia välilyöntejä. Poistimme ne Notepadin: Korvaa- toiminnolla. Poistimme lainausmerkit. Poistimme merkit, jotka vääristävät arvojen määrää (esim. jotkut lähtöasemat laskettiin kahdeksi, koska niiden nimi oli tarkennettu pilkulla) ja poistimme turhat väliviivat. Koska datan koko on oli niin suuri, täytyi se konvertoida arff- muotoon Wekassa online konverterin sijaan. Se löytyy alkunäytön tools- kohdasta ArffViewer. Puuttuvat arvot korvataan arvon keskiarvoilla.

4 Mallinnus

Ensin halusimme muodostaa kokonaiskuvan tilanteesta.

Kaupunkipyörillä ajettujen matkojen pituuden keskiarvo on n. 2,5 km ja keskimääräinen poljettu aika n. 14,1 min.

Rakennamme tyypilliset pyörien käyttötapaukset Klusteroinalla. Kluster moden arvoksi laitoimme "Use training set", klusterien määräksi laitoimme 10 ja Seediksi attribuuttien määrä 8. Huomasimme myös, että klusteroinnin tulos on ihan erilainen, jos laitamme Kluster moden arvoksi esim. Percentage split (66 %).

Tulos (Final cluster centroids):

[initial starting points (random):

```
Cluster 0: 2021-04-30T13:44:50,2021-04-30T13:54:11,81,Stenbäckinkatu,91,Seurasaari,90,2052
Cluster 1: 2021-04-29T18:33:19,2021-04-29T18:47:48,261,Asentajanpuisto,274,Voikukantie,597,3044
Cluster 2: 2021-04-11T16:11:03,2021-04-11T16:33:06,723,Säterinniitty,563,Mankkaanlaaksontie,1626,3808
Cluster 3: 2021-04-18T15:19:17,2021-04-18T15:19:42,127,Teurastamo,127,Teurastamo,1365,1
Cluster 4: 2021-04-12T14:05:15,2021-04-12T14:20:15,146,Pohjolanaukio,146,Pohjolanaukio,595,499
Cluster 5: 2021-04-10T22:23:30,2021-04-10T22:43:39,238,Pihlajamäki,140,Verkatehtaanpuisto,306,3859
Cluster 6: 2021-04-21T16:55:06,2021-04-21T17:15:50,40,Hakaniemi (M),141,Intiankatu,1004,4116
Cluster 7: 2021-04-19T21:53:55,2021-04-19T21:54:22,220,IdaAalbergintie,220,IdaAalbergintie,467,0
Cluster 8: 2021-04-22T18:18:38,2021-04-22T18:23:49,137,Arabiankauppakeskus,139,Muotoilijankatu,1472,1182
Cluster 9: 2021-04-26T18:05:54,2021-04-26T18:12:18,63,Jätkäsaarenlaituri,5,Sepänkatu,840,1194
```

| Final cluster centroids: | | Cluster# | |
|---------------------------------------|-------------------------|---------------------|---------------------|
| Attribute | Full Data (547336.0) | 0 (96192.0) | (13974.0) |
| inDeparture | 2021-04-21T14:17:52 | 2021-04-18T14:52:01 | 2021-04-30T17:03:3 |
| Return | 2021-04-14T19:26:09 | 2021-04-14T16:44:20 | 2021-04-26T18:48:3 |
| Departurestationid | 177.3063 | 53.3151 | 172.110 |
| Departurestationname | Itämerentori | Itämerentori | PetterWetterinti |
| Returnstationid | 176.538 | 51.7252 | 441.463 |
| Returnstationname | Itämerentori | Seurasaari | PetterWetterinti |
| Duration(sec.) | 906.3923 | 299.0677 | 1066.093 |
| Covereddistance(m) | 2484.612 | 2208.2456 | 3721.406 |
| < ----- > | | | |
| 2 (75536.0) | 3 (57586.0) | 4 (78544.0) | 5 (12102.0) |
| 2021-04-18T17:04:19 | 2021-04-29T18:30:42 | 2021-04-19T17:04:35 | 2021-04-19T21:06:35 |
| 2021-04-30T18:49:16 | 2021-04-30T18:33:14 | 2021-04-27T19:00:25 | 2021-04-30T18:11:57 |
| 601.4556 | 67.2507 | 122.0871 | 455.3576 |
| alto-yliopisto (M), Korkeakouluuaukio | Teurastamo | Pasilanasema | Merihaka |
| 601.4071 | 62.6992 | 124.464 | 104.9625 |
| alto-yliopisto (M), Korkeakouluuaukio | Teurastamo | Kalasatama (M) | Verkatehtaanpuisto |
| 881.6522 | 1534.4683 | 483.1836 | 1123.7305 |
| 1946.3406 | 3163.9895 | 2199.6364 | 3991.2251 |
| 6 (51002.0) | 7 (46000.0) | 8 (23706.0) | 9 (92694.0) |
| 2021-04-30T18:37:24 | 2021-04-19T15:43:49 | 2021-04-18T15:37:06 | 2021-04-30T22:29:31 |
| 2021-04-30T18:42:43 | 2021-04-30T19:10:25 | 2021-04-11T13:02:44 | 2021-04-30T23:03:39 |
| 119.4594 | 242.9575 | 137.5446 | 49.3994 |
| Hakaniemi (M) | Viikintiedepuisto | Arabiankauppakeskus | Itämerentori |
| 132.7256 | 241.346 | 140.6487 | 45.2135 |
| Intiankatu | Viikintiedepuisto | Kalasatama (M) | Itämerentori |
| 1176.9389 | 808.2614 | 2731.0772 | 905.9434 |
| 2771.3601 | 2199.3171 | 4980.8303 | 1991.7117 |

Clustered Instances

| | |
|---|--------------|
| 0 | 96192 (18%) |
| 1 | 13974 (3%) |
| 2 | 75536 (14%) |
| 3 | 57586 (11%) |
| 4 | 78544 (14%) |
| 5 | 12102 (2%) |
| 6 | 51002 (9%) |
| 7 | 46000 (8%) |
| 8 | 23706 (4%) |
| 9 | 92694 (17%) |

Esimerkki- stereotyppejä (rakennettu yllä olevista klustereista):

Klusteri 8: 45- vuotias Kirsi on ystävänsä kanssa tehdyltä shoppailu- reissulta kotiin

torstai- iltana kaupunkipyörällä palaava nainen. Mukaan löytyi mekko viikonlopuun juhlia varten.

Klusteri 4: 24- vuotias opiskelija Mauno on maanantai- iltapäivänä kaupunkipyörällä luennolta kahvilaan lukemaan siirtyvä mies.

Klusteri 7: 54- vuotias lähihoitaja Jaakko on palaamassa harjoittelupaikastaan iltavuorosta. Hänen suunnittelee ajavansa kaupunkipyörällä, mutta juuri kun hän on lähtemässä ajamaan, alkaa satamaan ja hän päättää käyttää vaihtoehtoista kulkuneuvoa.

Kaikista klustereista huomasimme ensimmäiseksi, että jo kymmenen klusterin joukossa on kaksi typpitapausta, joissa ei ole ajettu lainkaan, vaikka pyörä on otettu käyttöön vaan kaupunkipyörällä ajamisen ideasta on luovuttu tai sen käyttäminen ei ole onnistunut. Miksi?

Seuraavaksi huomasimme, että kohdassa 1 mainitsemassamme tavoitteessa lyhyistä pyöräilyajoista on onnistuttu. Tämä puhuu myös sen puolesta, että pyöräasemia pitää olla tarpeeksi tiuhaan. Ja myös huomasimme sen, että ajan mittaus ei aina toimi, koska ensimmäisen klusterin arvoksi tuli, että henkilö olisi ajanut 1.5 minuutissa reilun 2 kilometriä. Myös alkuperäisessä datassa oli pyöräilymatkoja, joita ei kannata ottaa huomioon, esim. pyörällä oli perutettu 9 m (ajomatka oli -9 m).

Pyörillä ajetaan todella paljon päivä- ja iltaaikaan klo 12 - 19. Tarkistimme asian rakentamalla 30 klusteria. Eli aamulla pyöräily ja varsinkin pyöräily öiseen aikaan on vähäistä verrattuna vaikkapa potkulautojen käyttöön, joissa öinen käyttö on suorastaan ongelmallista tapaturmien vuoksi (uutisissa on ollut). Joten alkoholin vaikutuksen alaisena ajamiseen ei tarvitse tällä hetkellä kaupunkipyöräilystä Helsingissä puhuttaessa kiinnittää huomiota, koska se tapahtuu eniten öiseen aikaan, jonka todistavat tapaturma- asemat, jotka hoitavat loukkaantuneita.

Toinen tärkeä asia, jonka halusimme selvittää on, paljonko eri kaupunkipyöräasemia on käytetty kyseisenä ajanjaksona. Sen saimme selville vain tutkimalla dataa. Vähiten kaupunkipyörä oli otettu käyttöön DerbyBusinessParkilta (92) ja Puutilankyläkartanolta (32). Myös pyörien palautuksia oli vähän Puutilankyläkartanolla (20), WorkShopHelsingillä (76) ja myös DerbyBusinessParkilla (102). Eniten pyöräilyjä on aloitettu Itämerentorilta (10 144) ja Ympyrätalolta (7074) ja eniten palautuksia on niille asemille on myös eniten palautuksia (Itämerentori 10 364).

5 Arviointi

Aamulla tapahtuvan pyöräilyn vähäisen määrän voisi ehkä selittää esimerkiksi sillä, että ihmiset eivät halua hikoilla ennen töitä, vaan valitsevat vaihtoehtoisen tavan mennä sinne ja tekevät paluumatkan kaupunkipyörällä. Öisen pyöräilyn puuttuminen (käytö esim. ravintolasta palatessa) johtuu kyseisenä tutkimusaikana vallinneesta koronatilanteesta. Se vaikuttaa myös aamuisen pyöräilyn vähyyteen, koska etätyö lisääntyi. Helsingin säästä kyseisen ajanjakson päivinä emme tässä tutkimuksessa ota huomioon tai sen vaikutusta pyöräilijämääriin, koska sitä emme voi ennustaa ja se ei vaikuta esim. pyöräasemien määrän suunnittelun samalla tavalla nyt.

6 Täytäntöönpano

Pyöräilymatkojen kestoaika on optimaalisen lyhyt eli sen takia ei asemia tarvitse lisätä. Mutta se on eri asia, pääsevätkö ihmiset niillä haluamiinsa paikkoihin tarpeeksi. Siitä voisi laatia kyselyn.

DerbyBusinessParkin ja Puutilankylänkartanon käyttömäärit ovat niin pieniä, sekä pyörän haltuunotto että palautusmäärit, että niiden asemien pois ottamista tai siirtämistä kannattaa harkita. Kannattaa harkita pyörien lisäämistä Itämerentorille, koska se on todella vilkas asema.

Vuonna 2019 HSL otti kaupunkipyörien rinnalle vuokrattavat sähköpotkulaudat.