

# Koululaisten liikkumisen edistäminen

Datapohjainen analyysi koulujen liikkumista edistävän toiminnan ja olosuhteiden yhteydestä koululaisten liikkumiseen ja kestävyyskuntoon

## Tiimi

Ville Meloni, kaupunginkanslia  
Nuutti Kytö, kaupunginkanslia  
Taija Savolainen, liikuntapalvelut  
Minna Paajanen, liikuntapalvelut  
Seppo Nyrkkö, Lohde Advance



*Kuva © Aki Rask/ Helsingin kaupunki*

# Sisällysluettelo

## Tiivistelmä

1. Kokeilun tausta
2. Kokeilun tavoitteet
3. Kokeilun keskeiset opit
4. Kokeilun eteneminen
5. Kokeilun tuotokset ja sisällölliset tulokset
6. Opit kokeiltavan ratkaisun tai toimintatavan mahdollisuuksista
7. Opit asiakkaiden tai palvelun käyttäjien tarpeista
8. Opit ratkaisun kehittämisestä teknisesti
9. Opit kokeilemisestä yleensä
10. Opit kokeiluprojektin arjen pyörittämisestä
11. Kokeilun tekninen ympäristö
12. Kokeilun data
13. Jatkopäätökset ja -ideat

# Tiivistelmä 1/2

## Kokeilun ydintavoitteet:

- **Selvittää**, onko kouluissa tehtävillä liikkumista edistävillä toimenpiteillä tai olosuhteilla yhteyttä koululaisten liikkumiseen tai kestävyyskuntoon
- **Kokeilla**, onko tämääntyyppinen data-analyysi järkevä keino yhdistää eri lähteistä kertyvää tietoa koululaisten liikkumista koskevien johtopäätösten tekemisen apuna

**Kokeilun nimi:** Datapohjainen analyysi koulujen liikkumista edistävän toiminnan ja olosuhteiden yhteydestä koululaisten liikkumiseen ja toimintakykyyn

**Uskomme, että** analysoimalla erilaisia kouluissa kerättäviä aineistoja yhdessä, voimme

- 1) selvittää, löytyykö kouluympäristössä selittäviä tekijöitä lasten ja nuorten liikkumiselle tai kestävyyskunnon ja näin tehdä toimenpidesuosituksia liikkumisen edistämiseksi,
- 2) havaita, mitkä kouluille suunnatut toimenpiteet tai aineistonkeruut ovat vähemmän relevantteja
- 3) ymmärtää, riittääkö kerätty tieto ja sen laatu vastaamaan tutkimuskysymyksiimme

**Kokeillaksemme tätä aiomme** analysoida useista eri tietolähteistä saatuja koulukohtaisia tietoja ja selvittää, mitkä kouluympäristöön ja koulun toimintaan liittyvät tekijät selittävät eniten lasten ja nuorten liikkumiskäyttäytymistä tai kestävyyskuntoa.

**Olemme oikeassa, jos** voimme tunnistaa

- 1) vaikuttavimpia toimenpiteitä tai tekijöitä koululaisten liikkumisessa ja kestävyyskunnossa
- 2) epärelevantteja tiedonkeruita tai toimenpiteitä
- 3) kerättävän tiedon laatutekijöitä, joita parantamalla tiedon hyödyntäminen mahdollistuu aiempaa paremmin

# Tiivistelmä 2/2

## Kokeilun opit:

- Analyysin avulla saimme tarkemman kuvan koulujen välisistä eroista koululaisten liikkumisessa ja kestävyyskunnossa sekä löysimme suuntaa-antavia tuloksia näiden yhteyksistä liikkumista edistävään toimintaan sekä muihin tarkasteltuihin muuttujiin, mutta datan laatu havaittiin heikoksi luotettavien johtopäätösten tekemiseksi

## Suositukset jatkotoimenpiteiksi:

- Jatkossa on tärkeä pohtia, miten koululaisten liikkumisesta tai toimintakyvystä kertyvää tietoa voidaan yhdistää tai hyödyntää tehokkaammin liikkumisen edistämisessä

**Kokeilun nimi:** Datapohjainen analyysi koulujen liikkumista edistävän toiminnan ja olosuhteiden yhteydestä koululaisten liikkumiseen ja toimintakykyyn

**Uskomme, että** analysoimalla erilaisia kouluissa kerättäviä aineistoja yhdessä, voimme

- 1) selvittää, löytyykö kouluympäristössä selittäviä tekijöitä lasten ja nuorten liikkumiselle tai kestävyyskunnolle ja näin tehdä toimenpidesuosituksia liikkumisen edistämiseksi,
- 2) havaita, mitkä kouluille suunnatut toimenpiteet tai aineistonkeruut ovat vähemmän relevantteja
- 3) ymmärtää, riittääkö kerätty tieto ja sen laatu vastaamaan tutkimuskysymyksiimme

**Kokeillaksemme tätä aiomme** analysoida useista eri tietolähteistä saatuja koulukohtaisia tietoja ja selvittää, mitkä kouluympäristöön ja koulun toimintaan liittyvät tekijät selittävät eniten lasten ja nuorten liikkumiskäyttäytymistä tai kestävyyskuntoa.

**Olemme oikeassa, jos** voimme tunnistaa

- 1) vaikuttavimpia toimenpiteitä tai tekijöitä koululaisten liikkumisessa ja kestävyyskunnossa
- 2) epärelevantteja tiedonkeruita tai toimenpiteitä
- 3) kerättävän tiedon laatutekijöitä, joita parantamalla tiedon hyödyntäminen mahdollistuu aiempaa paremmin

# 1. Kokeilun tausta

- **Ongelman kuvaus**

- Tietoa lasten ja nuorten liikkumisesta, terveydestä ja fyysisestä toimintakyvystä kerätään säännöllisesti Kouluterveyskyselyssä, LIITU-tutkimuksessa ja Move!-mittauksissa
- Myös koulujen liikkumista edistävää toimintaa sekä olosuhteita seurataan ja arvioidaan kansallisella Liikkuva koulu –nykytilan arvioinnissa, koulujen toimintasuunnitelmissa ja –kertomuksissa sekä kaupungin muissa tiedonkeruissa, kuten koulupihojen kuntoarvioissa tai harrastamiseen liittyvissä tiedonkeruissa.
- Merkittävä osa tiedosta on subjektiivista arviointia. Esimerkiksi koulun henkilöstö arvioi koulunsa liikkumista edistävän toimintansa laatua ja määrää, tai lapsi ja nuori arvioi omaa liikkumiskäyttäytymistään
- Nykymuodossaan tieto on kertynyt eritahtisesti eri järjestelmiin, raporteihin ja työyksiköihin. Erilaisten data-aineistojen ”ristiinajo” puuttuu. Tästä johtuen kokonaiskuvan muodostaminen lasten ja nuorten liikkumisen ilmiöstä on haasteellista ja työlästä. Toisin sanoen tietoa kerätään monin tavoin, mutta sitä ei tällä hetkellä kyetä läheskään täysimääräisesti hyödyntämään toiminnan kehittämisessä.
- Lisäksi Helsingissä kuntatasoinen raportointi vesittää alleen suuret erot lasten hyvinvoinnissa, kuten myös toiminnallisesti eri yksiköiden välillä

- **Rajaus**

- Kokeilu rajattiin koskemaan ainoastaan Helsingin kaupungin peruskouluja

## 2. Kokeilun tavoitteet: sisällölliset tavoitteet



### 1. Lasten ja nuorten liikkuminen ja kestävyyskunto kouluittain

- Miten lapset ja nuoret liikkuvat Helsingin kouluissa? Kuinka paljon on vähän liikkuvia lapsia? Kuinka moni täyttää liikkumisen suositukset?
- Mikä on lasten ja nuorten kestävyyskunnan taso Move!-mittausten perusteella Helsingin kouluissa?
- Millaisia eroja on sukupuolten, oppiasteiden ja koulujen välillä?



### 2. Kouluympäristön liikkumista edistävät olosuhteet ja toiminta

- Millaisia toimenpiteitä kouluissa tehdään liikkumisen edistämiseksi? Mitkä ovat yleisimpiä toimenpiteitä?
- Millaisia ovat Helsingin koulujen liikkumista tukevat olosuhteet (koulupihojen kunto, liikuntasalien koko, pyöräpysäköintipaikat)?



### 3. Kouluympäristön liikkumista edistävien olosuhteiden ja/tai toiminnan yhteydet koululaisten liikkumiseen ja/tai kestävyyskuntoon

- Ovatko kouluympäristön liikunnalliset olosuhteet tai kouluissa toteutetut liikkumista edistävät toimenpiteet yhteydessä koululaisten liikkumiseen tai kestävyyskuntoon? Onko sukupuolten ja luokka-asteiden välillä eroja?
- Voidaanko tunnistaa keskeisiä liikunnallisen koulun ominaispiirteitä? Voidaanko tehdä toimenpidesuosituksia kouluille?



### 4. Muiden selittävien tekijöiden yhteydet koululaisten liikkumiseen ja/tai kestävyyskuntoon

- Onko koulun sijainnin perusteella määritetty aluetason sosioekonominen summaindeksi yhteydessä lasten ja nuorten liikkumiseen tai fyysiseen toimintakykyyn?
- Onko koulussa järjestetyllä liikunnan painotetulla opetuksella yhteyttä näihin muuttujiin?

# Analyysiin sisällytetyt muuttujat, aineistot sekä aineistojen kerääjätaho



**Koulu keskiössä:**  
Kaikkia indikaattoreita tarkastellaan koulun tasolla, ei yksittäisten oppilaiden osalta



## Koulun liikkumista edistävä toiminta (selittävät muuttujat):

- Liikkuva koulu –toiminnan nykytilan arviointi (LIKES)
- Koulun toimintasuunnitelman kirjaukset liikunnallisen toimintakulttuurin edistämiseksi (Kasko)



## Koulun liikkumista edistävät olosuhteet (selittävät muuttujat):

- Koulupihan kuntoarvio (Stara)
- Koulun pyöräpysäköinnin nykytila (Kymp)
- Liikuntasalin koko suhteessa oppilasmäärään (Kuva)



## Oppilaiden liikkuminen ja toimintakyky (selitettävät muuttujat):

- Liikkuminen: kuinka monena päivänä liikkuu vähintään tunnin päivässä (THL Kouluterveyskysely)
- Kestävyyskunto: Move!-mittaustulokset kestävyystestistä (OPH)

## Muut selittävät tekijät:



- Alueen sosioekonominen summaindeksi (Kanslia)
- Liikunnan painotetun opetuksen järjestäminen (Kasko)

## 2. Kokeilun tavoitteet: tiedonkeruuseen ja tiedon hyödyntämiseen liittyvät tavoitteet

Koska kouluille suunnattujen tiedonkeruiden määrä on suuri, halusimme data-analyysin avulla myös:

- **Ymmärtää**, riittääkö annettu data ja sen tarkkuustaso vastaamaan tutkimuskysymyksiimme (esim. tarvittaisiinko tarkempaa, laadukkaampaa tai kattavampaa dataa)
- **Selvittää**, onko joistain tiedonkeruista mahdollista luopua tulevaisuudessa ja siten vähentää koulujen kuormitusta erilaisista tiedonkeruista

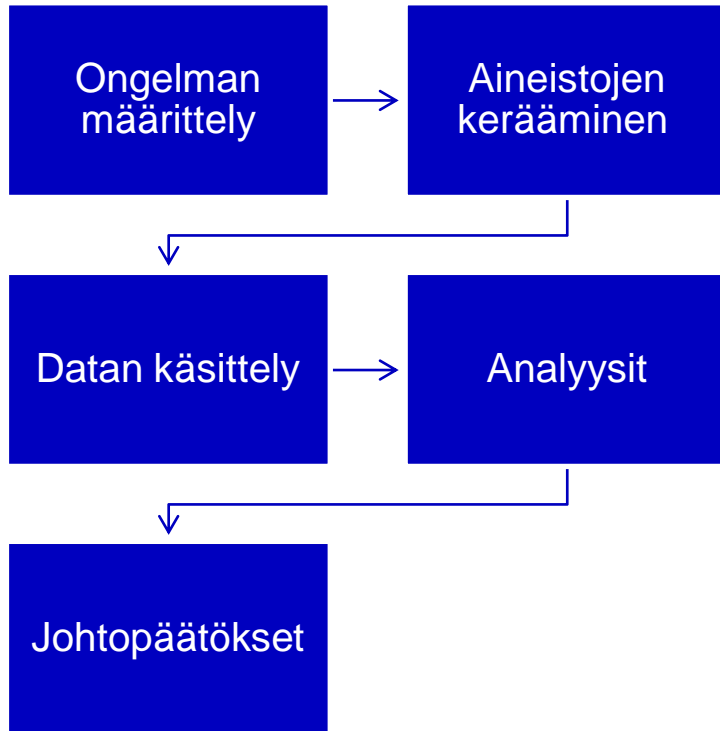
Lisäksi halusimme **kokeilla**, onko tämäntyyppinen data-analyysi järkevä keino yhdistää eri lähteistä kertyvää tietoa koululaisten liikkumista koskevien johtopäätösten tekemisen apuna.



# 3. Kokeilun keskeiset opit

- **Analyysin avulla** saimme tarkemman kuvan koulujen välisistä eroista koululaisten liikkumisessa ja kestävyyskunnossa, sekä löysimme suuntaa-antavia tuloksia näiden yhteyksistä liikkumista edistävään toimintaan sekä muihin tarkasteltuihin muuttujiin
- **Datan laatu jättää kuitenkin ilmaan hyvin paljon kysymysmerkkejä:** Valtaosa aineistoista perustuu joko lapsen tai nuoren, koulun henkilöstön tai muun asiantuntijan subjektiiviseen arviointiin. Erityisesti koulun liikkumista edistävää toimintaa kuvaava data ei kovinkaan hyvin sovellu tutkimuskäyttöön johtuen sen subjektiivisuudesta. Kaikkea dataa ei myöskään ollut saatavilla kaikista kouluista. **Kuinka vertailtavia tiedot ovat?**
- **Analyysi oli erittäin työläs toteuttaa** (aineistojen kerääminen eri tahoilta, datan siivous ja yhtenäistäminen, analysointi, tulosten tulkinta sekä johtopäätökset veivät runsaasti aikaa). Miten jatkossa tämäntyyppistä tietojen yhdistämistä ja analysointia on järkevä toteuttaa? Löydämmekö kestävämpiä ratkaisuja tiedon hyödyntämiseksi?
- **Tietojohtamisen työkalut ovat tällä hetkellä puutteellisia:** erilaisen datan säilytyspaikka, tietosuoja, datan omistajuus, analyysiosaaminen → mm. nämä asiat on ratkaistava, jos haluamme tulevaisuudessa tehdä vastaavanlaisia analyysyjä tai yhdistää eri tietolähteistä kertyvää dataa tehokkaammin
- Erilaisiin kyselyihin vastaaminen aiheuttaa kouluille paljon työtä. Jatkossa on pohdittava kriittisesti, **voidaanko osa tiedonkeruista yhdistää tai tiedonkeruuta toteuttaa muilla keinoin**, esim. liikkumisen edistämisen asiantuntijan havainnointina.

# 4. Kokeilun eteneminen



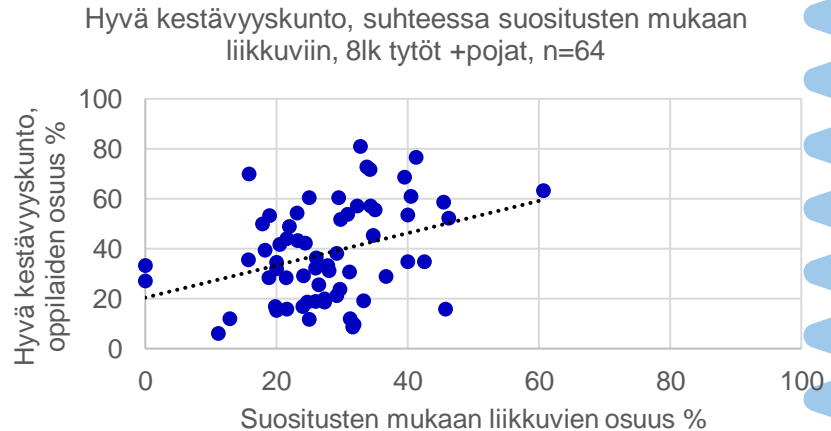
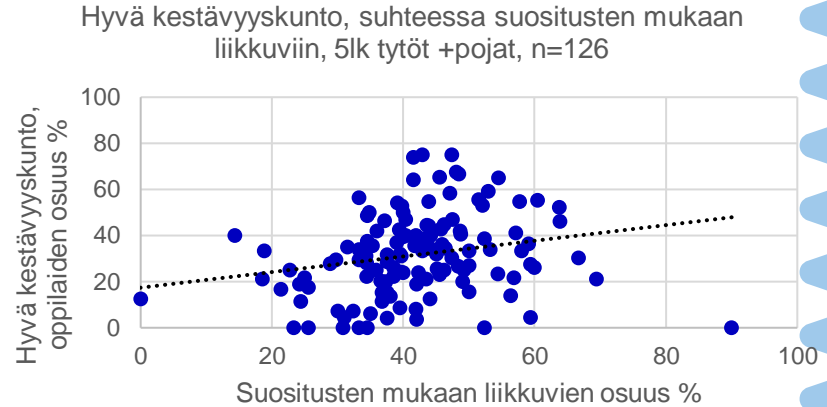
Ongelman määrittelyä lukuun ottamatta kaikki työvaiheet olivat työläämpiä ja hitaampia kuin alun perin ajattelimme.

Mikäli olisimme rajanneet tutkimuskysymysten tai aineistojen määrää pienemmäksi, olisi analyysin toteuttaminen ollut helpompaa.

Nyt käsitelimme noin kymmentä erilaista aineistoa, joiden kerääminen, yhdistäminen, käsittely, analysointi, tulosten visualisointi, tulkinta ja johtopäätösten tekeminen vei runsaasti aikaa.

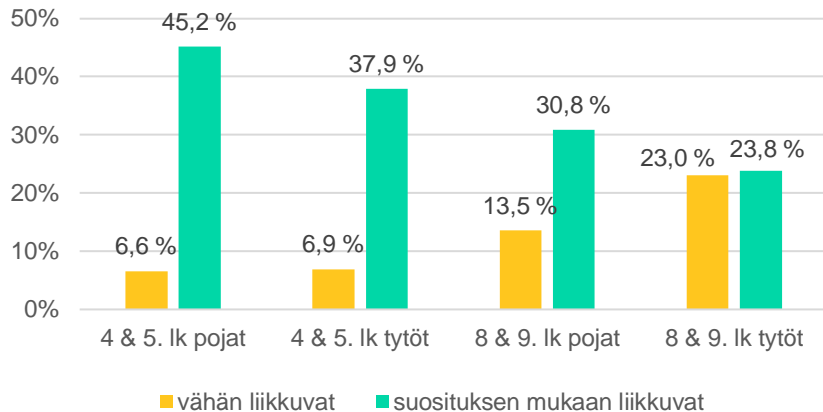
# 5. Kokeilun tuotokset teknisestä näkökulmasta

- Kokeilussa syntyi toistettava analyysiputki, joka vastaavalla datalla tuottaa määritetyt graafitiedot ja laskee korrelaatioiden suunnan ja tilastollisen merkitsevyyden.
- Tilastollinen korrelaatiomitta tukee hypoteeseja, kuten oppilaiden säännöllisen liikkumisen lisääminen on yhteydessä hyvään kestävyyskuntoon.
- Kokeilussa määritettiin joukko kuvailevia ja analysoivia arvoja, joiden kehityksen seuraaminen vuosittain voi kertoa toimenpiteiden vaikutuksesta lasten liikkumiseen



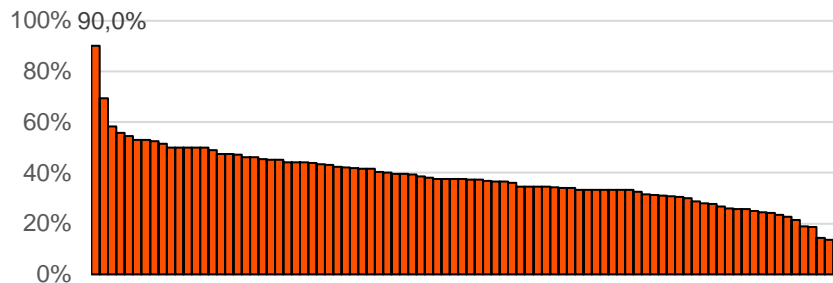
# 5. Tulokset: liikkuminen ja kestävyyskunto

Valtaosa koululaisista liikkuu liian vähän. Liikkuminen vähenee iän myötä erityisesti tytöillä.

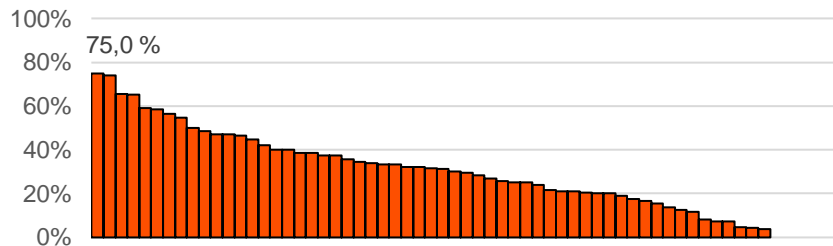


Koulujen väliset erot suosituksen mukaan liikkuvien, vähän liikkuvien sekä hyvän ja heikon kestävyyskunnon omaavien oppilaiden osuudessa olivat suuret kaikilla oppiasteilla ja sukupuolilla. Kuvioissa esimerkkinä alakouluikäiset tytöt. →

Vaihteluväli kouluittain liikkumisen suosituksen täyttymisessä 0-90 % 4 & 5. lk tytöillä (n=90)

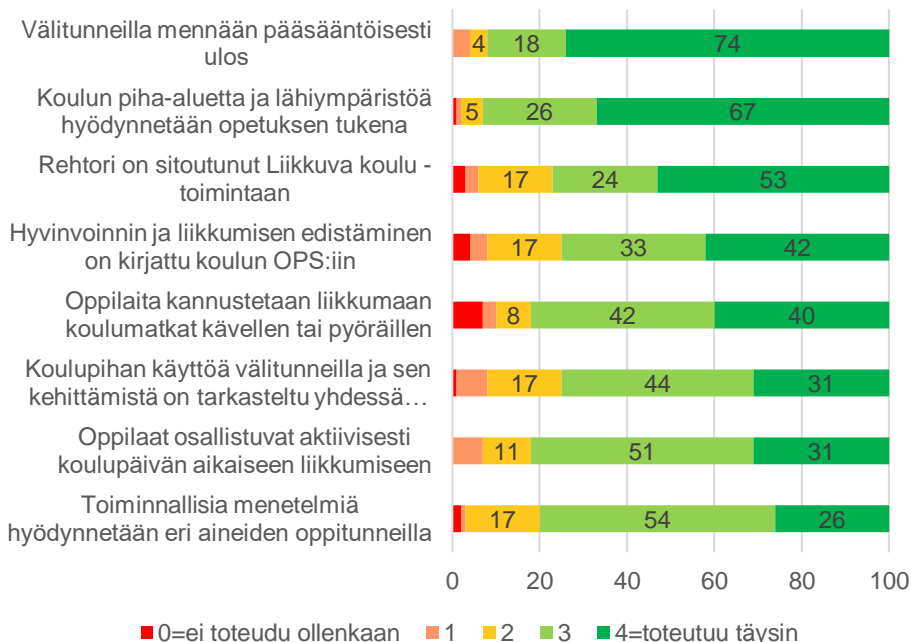


Vaihteluväli kouluittain hyvässä kestävyyskunnon 0-75 % 5. lk tytöillä (n=63)



# 5. Tulokset: liikkumista edistävä toiminta

Liikkuva koulu –nykytilan arviointi (n=75)



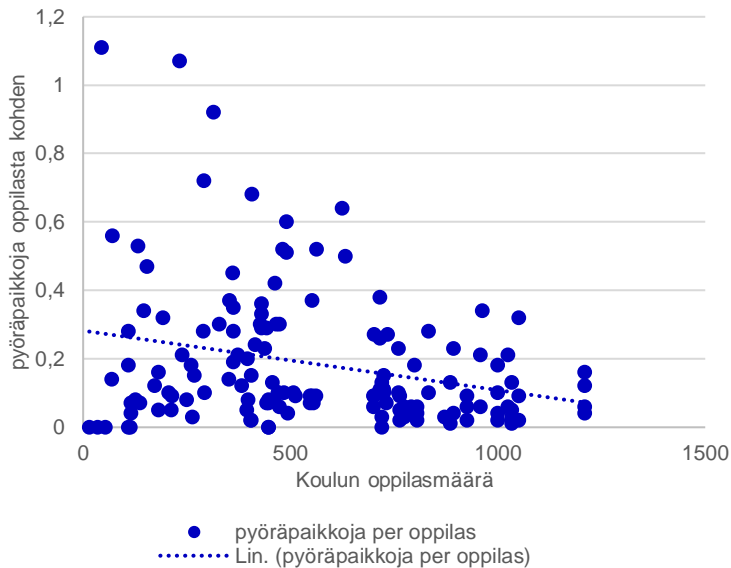
Liikkumista edistävät osa-alueet toimitasuunnitelmissa (n=102)



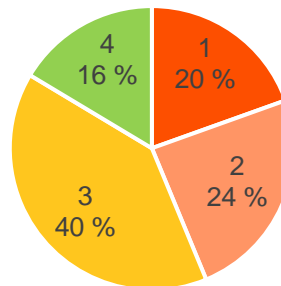
Lukuvuonna 2022-2023 Liikkuva koulu –nykytilan arvioinnin mukaan koulujen liikkumista edistävästä toimenpiteistä parhaiten toteutuvat välitunneilla ulkoilu sekä koulun pihan ja ympäristön hyödyntäminen opetuksessa. Oppilaita myös kannustetaan liikkumaan koulumatkat kävellen tai pyörällä yli 80 % vastanneista kouluista.

# 5. Tulokset: liikkumista edistävät olosuhteet

Pyöräpaikkoja oppilasta kohden on keskimäärin yksi viidelle oppilaalle (Hgin suositus on 1 paikka/2-3 oppilasta) Isoissa kouluissa pyöräpaikkoja on suhteessa vähemmän.

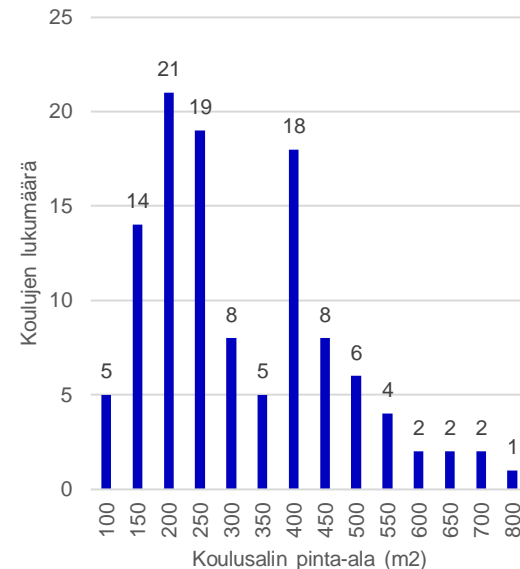


Lähes puolet koulupihoista on kunnostuksen tarpeessa Staran arvion mukaan



Arvio pihan kunnosta – jakauma kouluittain (1=kaipaa kunnostusta, 4=ei tarvetta korjata)

Valtaosa koulujen liikuntasaleista on alle 450 m<sup>2</sup>



Isoissa kouluissa pyöräpaikkoja on oppilasmäärään suhteutettuna vähemmän. Valtaosa koulujen liikuntasaleista on liian pieniä monipuoliseen harrastustoimintaan (alle 450 m<sup>2</sup>). Lähes puolet (44 %) koulupihoista on Staran arvion mukaan kunnostuksen tarpeessa. Toimenpiteitä koulujen liikkumista tukevien olosuhteiden parantamiseksi on jo tehty tai niitä on käynnissä.

# 5. Tulokset: liikkumista edistävän toiminnan yhteydet liikkumiseen ja kestävyyskuntoon

Suhteessa liikkumisen määrään ja hyvään kestävyyskuntoon positiivisina toiminnallisina asioina koko tutkimusjoukossa näyttäytyivät seuraavat asiat:

- Oppilaat osallistuvat aktiivisesti koulupäivän aikaiseen liikkumiseen
- Välitunneilla mennään pääsääntöisesti ulos
- Jokaiseen koulupäivään sisältyy vähintään 30 min liikkumisvälitunti
- Oppilaita kannustetaan liikkumaan koulumatkat kävellen tai pyöräillen

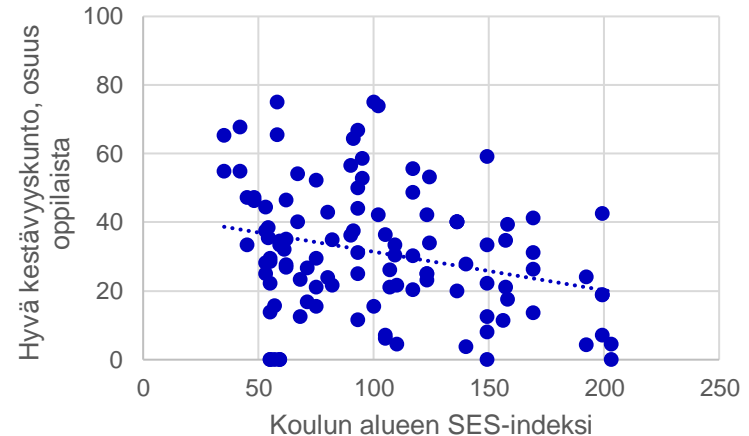
→ Huom! Analyysin tuloksista ei voida vetää johtopäätöksiä syy-seuraussuhteista.

Vaste	Muuttuja	r	p_arvo	otos_N
Suosituksen mukaan liikkuvien osuus	1.5 Liikkuva koulu -toiminta on kirjattu kuntastrategiaan, kunnan hyvinvointistrategiaan tai muuhun vastaavaan asiakirjaan.	-0,279	0,002	124
Suosituksen mukaan liikkuvien osuus	3.1 Oppilaat osallistuvat aktiivisesti koulupäivän aikaiseen liikkumiseen.	0,222	0,013	124
Suosituksen mukaan liikkuvien osuus	4.1 Välitunneilla mennään pääsääntöisesti ulos.	0,238	0,008	124
Suosituksen mukaan liikkuvien osuus	4.2 Jokaiseen koulupäivään sisältyy vähintään 30 min yhtenäinen liikkumisvälitunti.	0,279	0,002	124
Vähän liikkuvien osuus	1.3 Koulupäivän aikainen liikkuminen on kirjattu osaksi koulun lukuvuoden toimintasuunnitelmaa.	-0,231	0,010	124
Vähän liikkuvien osuus	3.1 Oppilaat osallistuvat aktiivisesti koulupäivän aikaiseen liikkumiseen.	-0,190	0,035	124
Vähän liikkuvien osuus	4.1 Välitunneilla mennään pääsääntöisesti ulos.	-0,179	0,047	124
Kestävyydeltään hyvässä kunnossa olevien osuus	5.1 Oppilaita kannustetaan liikkumaan koulumatkat kävellen tai pyöräillen.	0,292	0,001	124
Kestävyydeltään hyvässä kunnossa olevien osuus	2.1 Opetuksessa on sovittu käytännöistä, joilla istumisjaksot oppitunneilla katkaistaan.	-0,252	0,005	124
Kestävyydeltään hyvässä kunnossa olevien osuus	1.6 Rehtori on sitoutunut Liikkuva koulu -toimintaan.	-0,239	0,007	124
Kestävyydeltään hyvässä kunnossa olevien osuus	3.1 Oppilaat osallistuvat aktiivisesti koulupäivän aikaiseen liikkumiseen.	0,180	0,045	124
Kestävyydeltään heikossa kunnossa olevien osuus	5.1 Oppilaita kannustetaan liikkumaan koulumatkat kävellen tai pyöräillen.	-0,309	0,000	124
Kestävyydeltään heikossa kunnossa olevien osuus	3.1 Oppilaat osallistuvat aktiivisesti koulupäivän aikaiseen liikkumiseen.	-0,210	0,019	124
Kestävyydeltään heikossa kunnossa olevien osuus	1.6 Rehtori on sitoutunut Liikkuva koulu -toimintaan.	0,202	0,025	124
Kestävyydeltään heikossa kunnossa olevien osuus	5.3 Koululla on liikunnallista ja muuta aktiivista kerhotoimintaa koulupäivän yhteydessä.	0,179	0,047	124

# 5. Tulokset: muiden selittävien muuttujien yhteydet liikkumiseen ja kestävyyskuntoon

- **Alueen sosioekonomisella summaindeksillä ja 5. luokkalaisten kestävyyskunnolla sekä liikkumisella havaittiin tilastollisesti merkitseviä yhteyksiä.** Korkeamman SES-indeksin kouluissa heikko kestävyyskunto ja vähäisempi liikkuminen ovat vahvemmin edustettuina. Vastaavasti matalan SES-indeksin kouluissa hyvä kestävyyskunto ja suositusten mukaan liikkuminen ovat vahvemmin edustettuna. 8. luokkalaisilla vastaavia yhteyksiä ei havaittu.
- **Tilastollisia yhteyksiä koulujen olosuhteiden ja liikkumisen tai kestävyyskunnan välillä ei löytynyt.** Tämä voi johtua mm. siitä, että Staran kuntoarviossa ei arvioida koulupihan liikunnallisuutta, vaan välineiden teknistä kuntoa.
- **Liikunnan painotettu opetus** oli yhteydessä oppilaiden hyvään kestävyyskuntoon. Liikkumisen osalta yhteyksiä ei havaittu.

Hyvän kestävyyskunnan osuus suhteessa SES-indeksiin, 5. luokkalaiset (n=109)





# 6. Opit kokeiltavan ratkaisun tai toimintatavan mahdollisuuksista

Kokeilu on vahvasti data-analyttinen. Datatyöskentely jakautui kahteen osaan:

- Data-aineistoon tutustuminen ja muuttujien yhdenmukaistaminen analyysia varten.
- Selittävien ja kuvaavien muuttujien esittely (kuvailevat havainnot) sekä näiden väliset korrelaatiot (hypoteesien vahvistaminen)

Datan siivouksen ja yhdistämisen välineenä oli Python (Notebook) ja tiedon visualisoinnissa Excelin kaaviotyökalut. Eri muotoisten taulukoiden rakenteen saa Notebookeissa yhdistettyä raportointiystävälliseksi pienellä vaivalla. Raportointivälineenä ad hoc –tyyppiseen kehittämiseen Excel on ketterä (ns. ensimmäisen kierroksen prototyyppien kehitys). Kokeilun aikana saimme erityisesti oppeja datan esikäsittelyyn ja pyörittämiseen analyysien avulla.

**Tämäntyyppinen data-analyysi ei kuitenkaan ole tämän kokeilun perusteella tehokkain tapa yhdistää tai analysoida liikkumiseen liittyviä aineistoja.**

# 7. Opit asiakkaiden tai palvelun käyttäjien tarpeista

- Tiedonkeruiden määrän sijaan olennaista on jatkossa keskittyä käytössä olevan datan ja mittareiden laatuun, niiden jatkuvaan validointiin ja erilaisten raportointien johtopäätöksiin. Kokeilun ansiosta tunnistamme olemassa olevat tiedonkeruut aiempaa paremmin.
- Kokeilu teki uudesta, koulukohtaisesta tarkastelukulmasta käsin, näkyväksi merkittävät erot liikkumisessa ja fyysisessä toimintakyvyssä koulujen välillä.
- Kokeilu toteutettiin poikkileikkausasetelmassa. Olennaista jatkon kannalta on se, että lasten ja nuorten liikkumista sekä sitä edistävää toimintaa seurataan pitkäaikaisessa asetelmassa, mikä helpottaa myös eri asioiden välisten yhteyksien tunnistamista.
- Kokeilun asiakkaita voidaan katsoa olevan perusopetuspalvelun johto, asiantuntijat ja kehittäjät, kouluyhteisö kokonaisuudessaan, mutta lisäksi liikuntapalvelut, liikuntaa edistävät muut toimijat sekä liikkumisen edistäjät. Kokeilun tulosten ja havaintojen läpikäynti näiden toimijoiden kanssa on vielä kesken

# 8. Opit ratkaisun kehittämisestä teknisesti 1/3

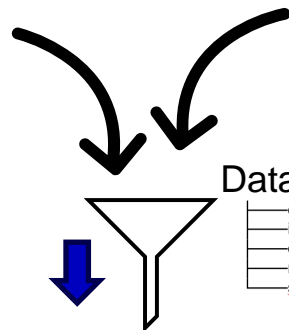
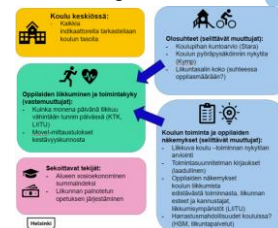
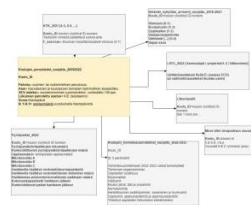
Kokeilussa lähdettiin liikkeelle olennaisesta kysymyksestä: Mitkä tekijät vaikuttavat lasten ja nuorten liikkumiseen/kestävyyskuntoon ja tämän pohjalta mitä taulukoita käytetään.

Tämän edetessä muodostui työnkulku:

- Datan siivousvaihe ja tiedon suojaus
- Lasketut kentät, Tietovarastomalli
- Poiminnat ja korrelaatiotestit tekevä Python-työkirja
- Analyysin tulosten asettelu graafeihin

Datan käsittelyputki toteutettiin Pythonin data- ja analyysikirjastojen avulla (Pandas, SciPy). Datan tarkastelun ja analyysiajojen työvälineeksi valittiin Jupyter Notebook, joka on Python-kielinen työkirjamuotoinen analyysiohjelma.

Lähdetieto:  
Kyselyjen Excel-tiedostot  
Move! PDF-raportit  
+ Tiedon suojaus

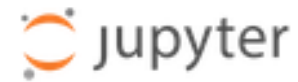


**Datan työstöalue**

- doc - documentation, reference
- input - place for RAW input files
- data - data files
- model - produced ML models
- src - Source code and notebooks

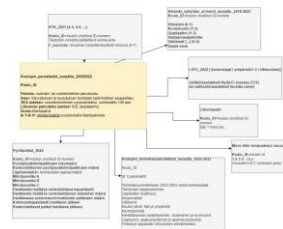
```
KTK Kouluterveyskysely, liikkuvuusryhmit
In [12]:
ksk = ksk_data[['ksk', 'i_2023']]
ksk
Out[12]:
   kouluk_01  vuosi  g_km_0230  g_km_240  g_km_260  g_km_280  g_km_300  kouluk_02  kouluk_03
0  101  2023    14.0    16.0    18.0    48.1    18.0    1  0
1  102  2023    2.0    16.0    44.2    37.2    1  0
2  103  2023    7.0    16.0    31.2    46.2    1  0
3  104  2023    2.0    4.1    24.5    68.4    1  0
4  105  2023    16.4    14.0    31.3    37.3    1  0
...
37  100  2023    9.5    25.0    30.1    34.2    0  0
38  101  2023    7.2    14.0    17.8    60.7    0  0
39  104  2023    71.6    38.8    32.3    28.0    0  0
40  200  2023    100.0    6.0    6.0    6.0    0  0
41  201  2023    12.8    33.3    31.9    32.2    0  0
```

Ryhmittely  
Ristiintaulukointi  
Analyysi



# 8. Opit ratkaisun kehittämistä teknisesti 2/3

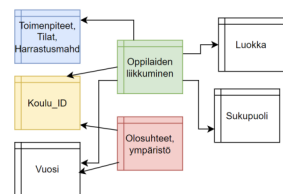
- Olenainen työvaihe on jakaa tiedot käsitelmallin mukaan mitattaviin faktatauluihin ja selittäviin dimensiotauluihin.
- Tämä mahdollistaa systemaattisen datan jakaumien analyysin ja korrelaatioiden laskennan kaikkien vastemuuttujien ja selittävien tekijöiden (toimenpiteet, olosuhteet) suhteen.
- Python-työkirja toimintatapana sallii nopean tilastomallien kehityksen.
- Graafeja varten tarvittava data tulostettiin Excel-tiedostoihin PowerPoint-raportointiin sopivaa muotoilua varten.



Lähdetaulut  
(Excel, PDF)



Käsitelmä, kyselyt,  
mittarit



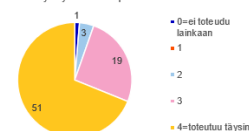
Tietovarastomalli

osajoukko	vaste	tekija	r	p_arvo	otos_N
yht	osuus_A	L_pyora_sailyty	-0.481639	0.000171	56
yht	osuus_A	L_valitunti_sali	0.297684	0.025867	56
yht	osuus_C	L_pyora_sailyty	0.435957	0.000784	56
yht	osuus_C	L_valitunti_sali	-0.271745	0.042766	56
yht	g_ex_012d	L_pyora_sailyty	-0.397438	0.002421	56

Tilasto ja korrelaatiomalli

## Työnkulku datasta tietoon

2.5 Koulun piha-alueita ja lähiympäristöä hyödynnetään opetuksen tukena.



# 8. Opit ratkaisun kehittämisestä teknisesti 3/3

Kokeilu sisälsi suurelta osin olennaiset osat tietovarasto- ja data-analyysi- ja raportointiprojektista.

Selkeä kansiomalli tiedostojen käsittelylle helpottaa ongelmien paikantamista ja tarvittaessa aineiston uudelleen käsittelyyn.

- **”input”** Erillinen kansio kerättävälle aineistolle
- **”data”** validointikansio (yhdistetyt ja ryhmitellyt kentät)
- **”src”** Poiminnat ja korrelaatiotestit tekevät Python-työkirjat ja apuohjelmat
- **”output”** Analyysin tulokset graafeihin omaan tulokansioon

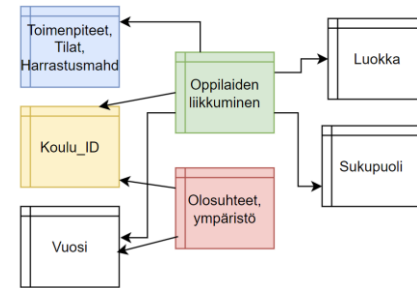
Työnkulun toistettavuuden välineet:

- Datakartta opastaa käyttäjää tiedostojen rakenteeseen
- Analyysitiedostojen ajojen automaatio toteutettu Jupyter Notebook-ohjelmissa

doc - documentation, reference  
input - **place for RAW input files**  
data - data files  
model - produced ML models  
src - Source code and notebooks



## Datan analysoinnin kansiorakenne



Datan rakenne kuvattuna datakartassa (muuttujat, taustatekijät)

# 9. Opit kokeilemisesta yleensä

- Näin laaja-alaisen ja monia aineistoja yhdistävän analyysin tulosten tulkinnassa ja johtopäätösten tekemisessä meni paljon odotettua enemmän aikaa.
- Myös aineistojen kerääminen eri tahoilta (THL, OPH, LIKES, kaupungin eri toimialat) oli työlästä ja aikaa vievää.
- Mikäli jatkossa samantyyppistä analysointia halutaan tehdä, tulisi teemaa tai tutkimuskysymyksiä rajata tarkemmin

# 10. Opit kokeiluprojektin arjen pyörittämisestä

- Kokeilulle on hyvä varata riittävästi työaikaa kaupungin kokeilutiimiltä. Myös kokeilun alussa on hyvä pohtia huolellisesti, mitä kaikkea kokeilun puitteissa on mahdollista toteuttaa.
- Konsulttityön osuudesta suuri osa kului aineiston käsittelyyn ja yhdistämiseen, sekä analysointiin. Tulosten visualisointiin, tulkintaan ja raportointiin olisi pitänyt resursoida enemmän aikaa
- Kokeilun raportointi viivästyi johtuen suuresta määrästä analyysejä, ja tulosten tulkinta vei runsaasti asiantuntijoiden aikaa

# 11. Kokeilun tekninen ympäristö

- Teknisinä ratkaisuinä käytettiin Azure ML ja Python notebookit:
  - Erittäin visuaalinen
- Azure, Document Intelligence – helppo ratkaisu kun tuotava aineisto on pdf:nä, tarvitsee poimintoja sääntöjen mukaan
- Muita Azure-kokeiluja joita olisi voinut kokeilla:
  - Keyword Extraction, Azure AI palvelut



# 12. Kokeilun data

- Kaupungin ulkopuolisilta toimijoilta kerättävän aineiston yhtenäistäminen analysoitavaan muotoon oli aikaa vievä prosessi
- Datan esikäsittely ja käsittely kannattaa tehdä ”kansioportaittain”:
  - Tietojen suojaus (koulujen nimet)
  - Sisään luettava data kannattaa käsitellä valmiiksi ohjelmallisesti erilliseen kansioon virheiden paikantamisen auttamiseksi
  - Työkirjamallia varten tarvitaan yhtenäiset sarakenimet, kuten Koulu\_ID, luokka, sukupuoli, vuosi
  - Raportointia varten tuotetut datat ja taulukot tulisi saada mahdollisimman lähelle käytettävää muotoa, kuten excelin scatter- tai bar plot –näkymään kelpaavaa muotoa.

# 13. Jatkopäätökset ja -ideat

- Jatkossa on tärkeä pohtia, miten koululaisten liikkumisesta tai toimintakyvystä kertyvää tietoa voidaan yhdistää tai hyödyntää tehokkaammin liikkumisen edistämisessä ja erilaisten toimenpiteiden kohdentamisessa. Vuoropuhelua on käytävä Kulttuurin ja vapaa-ajan sekä Kasvatuksen ja koulutuksen toimialojen välillä
- Analyysin tuloksia tulkitessa myös pohdittiin, voisiko jatkossa liikkumisen edistämiseen liittyvää tiedonkeruuta toteuttaa asiantuntijan havainnointina. Havainnointia voitaisiin toteuttaa esimerkiksi sellaisissa kouluissa, joissa koululaiset liikkuvat keskimääräistä enemmän, tai joissa liikkumisen edistämisen toimenpiteitä toteutetaan laajalla skaalalla. Näin voitaisiin havaita hyviä käytäntöjä, joita voitaisiin edistää erityisesti sellaisissa kouluissa, joissa liikkuminen on vähäistä.

# Kiitos!



*Kuva © Aki Rask/ Helsingin kaupunki*