

[Kurssin suoritus](#)

[Materiaalia](#)

[Ensimmäinen tapaaminen](#)

[Toinen tapaaminen](#)

[Kolmas tapaaminen](#)

[Neljäs tapaaminen \(6.10.\)](#)

[Viides tapaaminen \(13.10. ?\)](#)

[Kuudes tapaaminen \(20.10. ?\)](#)

## Kurssin suoritus

Kurssin laajuus on neljä opintopistettä ja se arvioidaan asteikolla hyväksyty/hylätty. Kurssin hyväksyty suoritus vaatii vähintään 60 % suorittautumista harjoitustehtävistä.

Kurssi käsittelee ohjelmoinnin perusteita ja erityisesti yhteiskuntatiedettä sovellussalueena. Tavoitteena on, että kurssin päätyttyä opiskelija ymmärtää minkälaisissa tutkimusongelmissa tietokoneohjelmista on hyötyä ja pystyy toteuttamaan itsenäisesti algoritmin (yksinkertaisen) tutkimuskysymyksen ratkaisemiseksi. Kurssin keskeisiä käsitteitä ovat

- muuttuja
- kontrollirakenteet (ehto- ja toistorakenteet)
- funktiot
- kirjastot

Työvälineinä käytetään sekä Scratch-ympäristöä että Python-ohjelmointikieltä.

## Materiaalia

Scratch

<http://bjc.berkeley.edu/website/curriculum.html>

<http://avoinoppikirja.fi/tite-scratch>

Python

[https://noppa.lut.fi/noppa/opintojakso/ct60a0200/materiaali/ohjelmointi\\_opas\\_v1.1.pdf](https://noppa.lut.fi/noppa/opintojakso/ct60a0200/materiaali/ohjelmointi_opas_v1.1.pdf)

(päivitetty linkki, sivunumerot tämän perusteella alla)

[http://www.ohjelmointiputka.net/oppaat/opas.php?tunnus=python\\_01](http://www.ohjelmointiputka.net/oppaat/opas.php?tunnus=python_01)

<http://solmu.math.helsinki.fi/2004/1/python.pdf>

<http://users.tkk.fi/ttsirkia/Python.pdf>

# Ensimmäinen tapaaminen

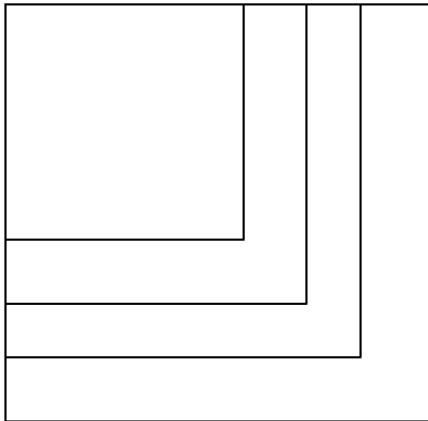
Luettava artikkeli keskustelua varten:

[https://www.dropbox.com/s/twd2tujklo50yqf/vajaa\\_alymysto.pdf?dl=0](https://www.dropbox.com/s/twd2tujklo50yqf/vajaa_alymysto.pdf?dl=0)

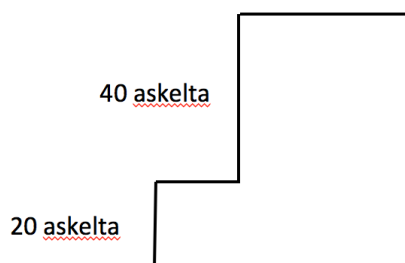
Tehtävät (käytä Scratchiä, <http://scratch.mit.edu/projects/editor/>)

Käytä mahdollisimman vähän komentoja!

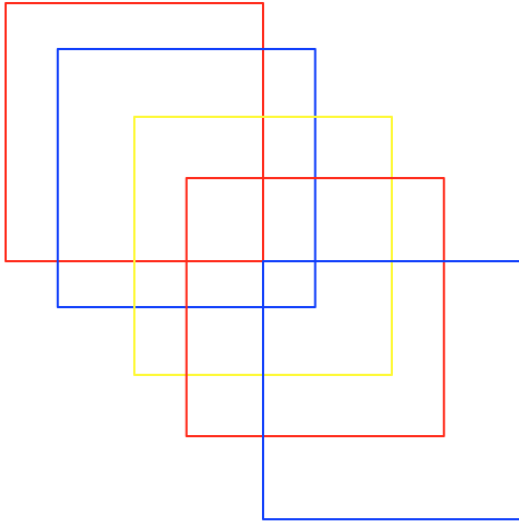
1. Piirrä neliö
2. Piirrä kolmio
3. Piirrä viisikulmio
4. Piirrä porraskuvio
5. Piirrä (mahdollisimman tarkka) ympyrä
  
6. Piirrä väriä vaihtava viiva
7. Piirrä katkoviiva
8. Piirrä (käytä muuttujaa, jos voit!)



9. Piirrä (käytä muuttujaa, jos voit)



10. Piirrä neliö, jonka pystysivut ovat punaiset ja vaakasisivut keltaiset
11. Piirrä kuvio 7, mutta joka toinen neliö keltaisella viivalla (vinkki: jakojäännös)
12. Piirrä



## Toinen tapaaminen

**Luettava artikkeli** keskustelua varten:

Lazer et al. (2009): Computational Social Science

[https://www.dropbox.com/s/hdc4a41j14ed3fn/lazer\\_computational.pdf?dl=0](https://www.dropbox.com/s/hdc4a41j14ed3fn/lazer_computational.pdf?dl=0)

### **Syventäviä artikkeleita**

Raento, Mika, Antti Oulasvirta, and Nathan Eagle. 2009. "Smartphones. An Emerging Tool for Social Scientists." *Sociological Methods & Research* 37 (3): 426–454.

Adamic, Lada A, and Natalie Glance. 2005. "The Political Blogosphere and the 2004 U.S. Election: Divided They Blog." In *Proceedings of the 3rd International Workshop on Link Discovery*, 36–43.

Eagle, N., & Pentland, A. 2006. Reality mining: sensing complex social systems. *Personal and ubiquitous computing*, 10(4), 255-268.

**Tehtävät** (käytä Scratchiä, <http://scratch.mit.edu/projects/editor/>)

Käytä mahdollisimman vähän komentoja!

1. Laske Scratchillä 5+7 ja sijoita tämä uuteen muuttujaan "summa"
2. Laske Scratchillä 10-3 ja sijoita tämä uuteen muuttujaan "erotus"
3. Laske  $1+2+3+4+5+6+7+8+9+10$  (vinkki: käytä toistoa)

4. Laske  $5*5*5*5*5$
5. Laske  $1*2*3*4*5*6$
  
6. Tarkista, onko 5 pienempi kuin 7; jos on kissan pitää sanoa "Pienempi"
7. Tarkista, onko 44 jaollinen luvulla 3; jos on kissan pitää sanoa "On jaollinen"
8. Kysy käyttäjältä lukua ja tarkista, onko se jaollinen luvulla 8. Jos on, kissan pitää sanoa "On jaollinen"
  
9. Laske yhteen kaikki kolmella jaolliset luvut välillä välillä 1-100
10. Laske yhteen kaikki kolmella tai viidellä jaolliset luvut välillä 1 - 100
  
11. Laske yhteen lukuja  $1+2+3+\dots$  kunnes summa on yli 157.
12. Pyydä käyttäjältä uusi luku, kunnes syötetään luku -1. Laske käyttäjien syöttämien lukujen summa.

## Kolmas tapaaminen

### Keskeisiä sekä hyödyllisiä käsitteitä

- muuttuja, operaattori
  - s. 5-7
  - s. 10-12
- ehtolause (if-lause)
  - s. 26-29
- toistolause, toistorakenne (for-lause)
  - s. 37
  - s. 39-40
- funktio
  - s. 43-45
  - s. 46-48
- sanojen käsittely (merkkijonot, string)
  - s. 16-20
- jaollisuus
  - s. 10-12

### Luettava artikkeli

Cioffi-Revilla, Claudio. 2010. "Computational Social Science." *Wiley Interdisciplinary Reviews: Computational Statistics* 2 (3) (May 12): 259–271.

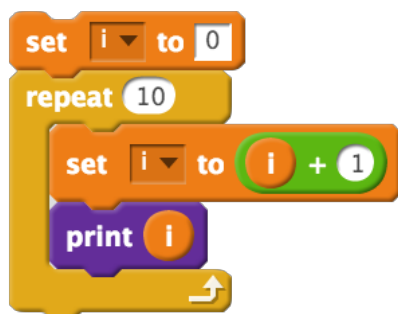
### Syventävät artikkelit

Squazzoni, F., W. Jager, and B. Edmonds. 2013. "Social Simulation in the Social Sciences: A Brief Overview." *Social Science Computer Review* 32 (3) (December 6): 279–294.

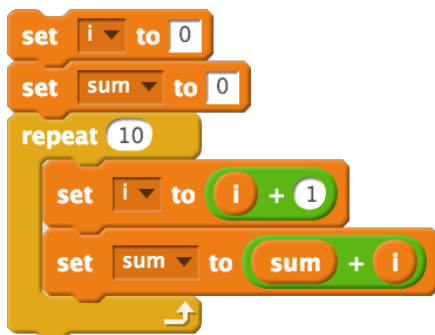
Brugger, N. 2013. "Historical Network Analysis of the Web." *Social Science Computer Review* 31 (3).

**Tehtävät** (käytä Pythonia, <http://humanisti.fixme.fi:8888/login?next=%2F> )

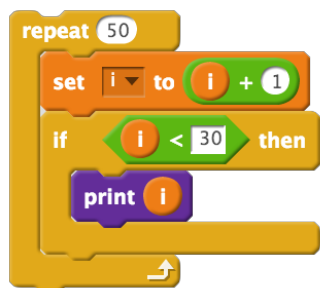
1. Toteuta Pythonilla seuraava Scratch-ohjelma



2. Toteuta Pythonilla seuraava Scratch-ohjelma (sum pitäisi olla 55)



3. Toteuta Pythonilla seuraava Scratch-ohjelma



+

4. Laske lukujen 1 ... 11 kertoma, eli  $1 \cdot 2 \cdot \dots \cdot 11$  (on 39916800)
5. Laske lukujen 1 ... 10000 neljällä jaollisten lukujen summa (on 12505000)  
Vihje: jakojäännös, esimerkiksi  $5 \% 3 \rightarrow 2$ ,  $10 \% 3 \rightarrow 1$ ,  $10 \% 5 \rightarrow 0$
6. Laske, kuinka moni luku välillä 1 ... 100 on jaollinen kolmella (on 33)

7. Toteuta funktio summa(yläraja, jaettava), joka laskee summan yläraja, ja laskee yhteen luvut jotka ovat jaollisia jaettavalla
8. Käyttäen summa-funktioa, laske yhteen kaikki 2, 3, 4, ... 10 jaolliset luvut välillä 1-100
9. Käyttäen summa-funktioa, selvitä millä jaettavalla 2, 3, 4, ... 10 summa on suurin luvuilla välillä 1-50.
10. Toteuta funktio verot( palkka ), joka saa syötteenä henkilön tuloverotuksessa verotettavan palkkasumman ja tämän perusteella laskee maksettavan veron määrän. Yksinkertaisuuden takia, älä mieti verovähennyksiä. Tuloverotaulukko [http://vero.fi/fi-FI/Syventavat\\_veroohjeet/Henkiloasiakkaan\\_tuloverotus/Valtion\\_tuloveroasteikko\\_2014\(30521\)](http://vero.fi/fi-FI/Syventavat_veroohjeet/Henkiloasiakkaan_tuloverotus/Valtion_tuloveroasteikko_2014(30521))
11. Muuttujiin ei tarvitse säilöä vain numeroita. Kokeile määritellä uusi muuttuja ja laita sanasi lainausmerkkien väliin  
sana = "valitsemasi sana"  
Selvennä itsellesi, mitä seuraavat asiat tekevät  
sana + "!"  
"Hän sanoi " + sana  
"a" in sana  
"ö" in sana
12. Tee funktio vokaali joka saa parametrina sanan ja tarkistaa onko sanassa vokaalia. Jos on, funktio palauttaa arvon True, muutoin se palauttaa arvon False
13. Tee funktio tulot( tukikuukaudet ), joka laskee kuinka monta euroa saat ansaita vapaita tukia jos saat opintotukea tukikuukaudet verran. Katso [http://www.kela.fi/tulot\\_opiskelijan-omat-tulot](http://www.kela.fi/tulot_opiskelijan-omat-tulot)
14. Tee funktio, joka tarkistaa onko luku alkuluku. Alkuluku on luku, joka on jaollinen vain yhdellä ja luvulla itsellään.
15. Kuinka monta alkulukua on välillä 2-150?
16. Tulosta luvut 1-100 seuraavasti: jos luku on jaollinen kolmella, tulosta Fizz, jos luku on jaollinen viidellä, tulosta Buzz, jos luku on jaollinen sekä viidellä että kolmella, tulosta FizzBuzz; muutoin tulosta luku
17. Tee funktio, joka tarkastaa onko syötteenä olevassa tekstissä sanaa "Python" - jos on, palauta True, muutoin palauta False
18. Toteuta funktio merkki( maara, merkki), joka palauttaa maaran verran merkkiä merkki. Esimerkiksi merkki( 5, 'A' ) palauttaa AAAAA
19. Toteuta funktio kolmio( korkeus ), joka tulostaa korkeus-muuttujan kokoisen kolmion, esimerkiksi kolmio( 4 ) tulostaa seuraavan kuvion  
\*  
\* \* \*  
\* \* \* \* \*  
\* \* \* \* \* \* \*
20. Mikä on lausekkeen  $2x^3 + 4x^2 + 6x + 8$  suurin arvo välillä -100, 100 (vain kokonaisluvuilla)
21. Mikä on lausekkeen  $7x^3 + 2x^2$  pienin arvo välillä -135, 135
22. Millä arvolla lausekkeen  $5x^4 + 8x$  arvo on suurin välillä -150, 150?

## Neljäs tapaaminen (6.10.)

### Keskeisiä ja hyödyllisiä käsitteitä

- lista (list, array)
  - s. 65-69
- ehtolause (if-lause)
  - s. 26-29
- toistolause, toistorakenne (for-lause)
  - s. 37
  - s. 39-40
- funktio
  - s. 43-45
  - s. 46-48
- sanakirja (dictionary)
  - s. 77-78

### Artikkeleita (valitse yksi, tai ehdota omaasi)

Saunders-Newton, D., and H. Scott. 2001. "But the Computer Said!": Credible Uses of Computational Modeling in Public Sector Decision Making." *Social Science Computer Review* 19: 47–65.

Jansson, Gabriella, and Gissur Ó Erlingsson. 2014. "More E-Government, Less Street-Level Bureaucracy? On Legitimacy and the Human Side of Public Administration." *Journal of Information Technology & Politics* 11 (3) (August 3): 291–308.

Boyd, Danah, and Kate Crawford. 2012. "Critical Questions for Big Data." *Information, Communication & Society* 15 (5) (June): 662–679.

### Tehtävät

1. Toteuta funktio pienin( a, b, c ), joka palauttaa pienimmän luvuista a, b, c
2. Tee lista jossa on luvut 0, 5, 7, 10, 8, 3
3. Listaa edellisen taulukon alkiot
4. Laske edellisen taulukon alkioiden summa
5. Mikä on taulukossa olevista luvuista suurin?
6. Lisää listaan kaikki kolmella jaolliset luvut 1-250 välillä
7. Toteuta funktio lista( ala, ylä, jaollisuus ), mikä palauttaa uuden listan jossa on kaikki luvut ala ja ylä välillä ja jotka ovat jaollisia luvulla jaollisuus

8. Toteuta funktio summa( lista ), joka laskee listan sisällön summan
9. Toteuta funktio keskiarvo( lista ), joka laskee listan sisällön keskiarvon
10. Toteuta funktio varianssi( lista ), joka laskee listan sisällön varianssin
11. Toteuta funktio suurin( lista ), joka palauttaa listan suurimman alkion
12. Toteuta funktio kaanna( lista ), joka palauttaa listnan käänteisessä järjestyksessä
13. Toteuta sanakirja, jossa on on nimesi ja sähköpostiosoitteesi
14. Toteuta sanakirja, jossa on eduskunnan puolueet (avaimina) ja niiden paikkamäärä eduskunnassa (arvoina)
15. Toteuta sanakirja jonka avaimena on postinumero ja arvona kaupunginosan nimi. Lisää muutama esimerkkikaupunginosa. Testaa tulostamalla postinumeroita ja niitä vastaavia kaupunkeja!
16. Kirjoittamalla from sanakirja import \* (omalle rivilleen) saat käyttöön sanakirjan "kaupungit" (eli, voit käyttää sitä muuttujana) joka sisältää maita ja niitä vastaavia kaupunkeja. Tarkista
  - a. onko Ruotsi kaupungit-sanakirjassa
  - b. mikä kaupunki Suomesta on sanakirjassa?
17. Tee uusi sanakirja, jonne lisäät luvut 1-15 (avaimina) ja näiden lukujen neliöt (luku kerrottuna itsellään) arvoina
18. Tee uusi sanakirja, jonne lisäät luvut -20 - 20 (avaimina) ja avaimina kunkin luvun lisättynä viidellä
19. Kirjoittamalla from tulokset import \* (omalle rivilleen) saat käyttöön listan vaalit, joka sisältää puolueiden nimiä, esimerkiksi "KESK", "SDP", ... Laske kunkin puolueen paikkamäärät tässä listassa.
20. Tee funktio kirjaimet( sana ), joka palauttaa sanakirjan sanassa olevista kirjaimista ja niiden määristä. Esimerkiksi kissa palauttaisi { 'k' : 1, 'i': 1, 's' : 2 , 'a' : 1 }
21. Samoin löytyy lista numerot - laske kuinka monta kertaa eri numerot numero esiintyvät listalla. Esimerkiksi syötteellä [5, 4, 5] tuloksena on { 5 : 2, 4 : 1 }
22. Myös löytyy avaimia-niminen sanakirja. Listaa kaikki tämän sanakirjan avaimet ja niitä vastaavat arvot

## Viides tapaaminen (13.10. ?)

### Keskeisiä sekä hyödyllisiä käsitteitä

- tiedoston lukeminen
  - s. 52-53
  - s. 56
- toistolause, toistorakenne (for-lause)
  - s. 37
  - s. 39-40
- ehtolause (if-lause)
  - s. 26-29
- sanojen käsittely (merkkijonot, string)



- s. 16-20
- s. 58-61
- muuttuja, operaattori
  - s. 5-7
  - s. 10-12

Katso esimerkkejä tiedostojen käsittelyyn [tästä](#).

### **Luettava artikkeli**

Ukkonen, Antti. 2013. “‘Big Data’ Ja Laskennalliset Menetelmät.” In *Otteita Verkosta: Verkon Ja Sosiaalisen Median Tutkimusmenetelmät*, edited by Salla-Maaria Laaksonen,, Janne Matikainen, and Minna Tikka. Tampere: Vastapaino.

### **Syventävät artikkelit**

Boyd, Danah, and Kate Crawford. 2012. “Critical Questions for Big Data.” *Information, Communication & Society* 15 (5) (June): 662–679.

Jurek, Steven J., and Anthony Scime. 2013. “Achieving Democratic Leadership: A Data-Mined Prescription.” *Social Science Quarterly*

Tumasjan, a., T. O. Sprenger, P. G. Sandner, and I. M. Welp. 2011. “Election Forecasts With Twitter: How 140 Characters Reflect the Political Landscape.” *Social Science Computer Review* 29: 402–418. doi:10.1177/0894439310386557.

### **Tehtäviä**

1. Tiedostossa vaalit\_tweetit.txt on kaikki kuntavaalien aikana kirjoitetut tweetit, joissa on joko #vaalit, #kunnallisvaalit tai #kuntavaalit -hashtag.
  - a. Kuinka monta tweettiä on?
  - b. Kuinka monessa tweetissä mainitaan sana jytky?
  - c. Kuinka moni tweetti sisältää linkin?

Alkuun pääset seuraavalla koodilla

```
for rivi in open('vaalit_tweetit.txt', 'r'):
    rivi = rivi.strip() ## poistetaan turhat valit alusta ja lopusta
    print rivi
```

2. Tiedostossa vaalit\_verkot.csv on kaikki tweetit joissa mainitaan joku kunnallisvaaliehdokkaista tai jonka on kirjoittanut kunnallisvaaliehdokas, muodossa lähettäjä,vastaanottaja.

- a. Kuinka monta tweettiä on lähetetty?
- b. Kuinka monella tweetillä ei ole vastaanottajaa, vaan ne ovat “huutoja”?  
Tweetillä ei ole vastaanottajaa, jos vastaanottaja on 0.

Alkuun pääset koodilla

```
for line in open('vaalit_verkot.csv', 'r'):
    line = line.strip() ## poistetaan turhat valit alusta ja lopusta
    kenelta = line.split(',')[0] ## kaikki ennen pilkkua
    kenelle = line.split(',')[1] ## kaikki pilkun jälkeen
    print kenelta, '->', kenelle
```

3. Verottaja on luovuttanut [verotustietoja](#) avoimena datana. Tiedostossa verot\_2012.csv on vuoden 2012 yritysverojen maksuun liittyviä tietoja muodossa Verovuosi;Y-tunnus;Verovelvollisen nimi;Verotuskunta;Verotettava tulo;Maksuunpannut verot yhteensä;Ennakot yhteensä;Veronpalautus;Jäännösvero
  - a. Mikä yritys maksoi eniten veroja?
  - b. Mikä yritys sai eniten mätkyjä?
4. Kuka oli aktiivisin tweettaaja kunnallisvaalien alla? (Haastavempi lisätehtävä)
5. Kuinka monta prosenttia kymmenen eniten veroja maksavaa yritystä maksoivat veroista? (Haastavempi lisätehtävä)
6. Minkä puolueen ehdokkaat puhuvat eniten jytkestä? Tiedostossa politicians\_twitter.txt on jokaista Twitter-tunnusta vastaava luulemamme puolue ja politicians\_tweets.txt on tweettejä, alussa tweettaajan tunnus. (Haastavempi lisätehtävä)

## Kuudes tapaaminen (20.10. ?)

### Keskeisiä sekä hyödyllisiä käsitteitä

- tiedoston lukeminen
  - s. 52-53
  - s. 56
- toistolause, toistorakenne (for-lause)
  - s. 37
  - s. 39-40
- sanojen käsittely (merkkijonot, string)
  - s. 16-20
  - s. 58-61
- muuttuja, operaattori
  - s. 5-7
  - s. 10-12
- funktio

- s. 43-45
- s. 46-48
- ehtolause (if-lause)
  - s. 26-29

## Luettava artikkeli

Grimmer, Justin, and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21: 267–297.

## Syventäviä artikkeleita

Levy, K. E. C., and M. Franklin. 2013. "Driving Regulation: Using Topic Models to Examine Political Contention in the U.S. Trucking Industry." *Social Science Computer Review* (December 2).

Woolf, BP, TJ Murray, Xiaoxi Xu, LJ Osterweil, Lori Clarke, Leah Wing, and Ethan Katsh. 2012. "Computational Predictors in Online Social Deliberations." In *Fifth International AAAI Conference on Weblogs and Social Media*, 603–606.

Weber, Ingmar, Antti Ukkonen, and Aris Gionis. 2012. "Answers, Not Links." In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining - WSDM '12*, 613–622.

Shor, Eran, Arnout van De Rijdt, Charles Ward, Soussan Askar, and Steven Skiena. 2014. "Is There a Political Bias? A Computational Analysis of Female Subjects' Coverage in Liberal and Conservative Newspapers." *Social Science Quarterly* 95 (5).

## Tehtäviä

1. Tiedostossa tweets\_per\_party.txt on puolueita ja tweettejä. Mikä puolue puhuu eniten palveluista? Oletko yllättynyt?
2. Tiedostossa tweets\_per\_party\_testi.txt (käytä testitiedostoa, se on lyhyempi ja tulokset tulevat minuuteissa, ei tunneissa) on ehdokkaiden tweettejä ja puolueita; laske viestien positiivisuutta ja negatiivisuutta? Oliko SDP vai Kokoomus positiivisempi viesteissään? Käytä mieliala-funktiota avuksesi:

```
import requests
import re
import urllib
```

```
numeroerotin = re.compile(".*<b>([0-9])</b>.*<b>([+-]?[0-9])</b>.*")
```

```

## anna teksti, ja saat arvion siitä onko sen äänensävy positiivinen tai negatiivinen
## palauttaa luvun äänensävyistä, negatiiviset luvut ovat negatiivista mielialaa
## kokeile itse, http://sentistrength.wlv.ac.uk/results.php
def mieliala( text ):
    ## käydään internetissä nettipalvelussa arvioimassa tätä
    data = urllib.urlencode( { 'text' : text, 'submit' : 'Detect+Sentiment+in+Finnish' } )
    url = 'http://sentistrength.wlv.ac.uk/results.php?' + data
    sivu = requests.get( url )
    ## luetaan nettisivulta sekä positiivinen että negatiivinen fiilis
    numerot = numeroerotin.search( sivu.text )
    positiivinen = int( numerot.group(1) ) - 1
    negatiivinen = int( numerot.group(2) ) + 1
    ## päätellään oliko tänä viesti nyt positiivinen vai negatiivinen
    if -1.5 * negatiivinen > positiivinen:
        return negatiivinen
    else:
        return positiivinen

```

3. Vertaile vihreiden ja vasemmiston tweeteissä olevia sanoja, löydätkö näiden puolueiden välillä eroja - esimerkiksi katso yleisintä 50 sanaa? tweets\_per\_party.txt on kaikki puolueiden tweetit samalla tavalla kuin edellisissä tehtävissä.
4. Kerää ohjelmallisesti ensimmäisen sivun ehdotusten nimet Otakantaa.fi:stä sääntöjen purkamisen neuvotteluista ( [https://www.otakantaa.fi/fi-FI/Selaa\\_hankkeita/Parempaa\\_saantelya/Vahemman\\_ja\\_parempaa\\_saantelya\\_miten\\_Mi\(28294\)?page=1](https://www.otakantaa.fi/fi-FI/Selaa_hankkeita/Parempaa_saantelya/Vahemman_ja_parempaa_saantelya_miten_Mi(28294)?page=1) ). Voit käyttää apunasi yllä mainittua requests -kirjastoa sekä [BeautifulSoapia](#).
5. Kerää tieto ensimmäisen sivun lisäksi kaikilta sivuilta.
6. Käy hakemassa nimien lisäksi varsinainen kannanottoteksti ja tulosta se otsikon viereen