

Työkalut tutuksi

Avoin data opetuksessa -koulutus

Maanantai 12.4.2021

Päivän tavoitteet

- Tutustutaan koulutuksessa käytettäviin työkaluihin - Jupyter Notebookiin & Python-ohjelmointiin
 - Jupyter Notebookin perustoiminnot (solujen luominen, solutyypit, koodin suorittaminen)
 - Data-analyysin kannalta tärkeimmät Python-paketit ja niiden käyttö
 - Käydään läpi esimerkki, jonka jälkeen päästään itse koodaamaan!

Jupyter Notebook

- Interaktiivinen ympäristö/dokumentti, jossa voidaan yhdistää tekstiä, kuvia, videoita ja koodia
- Opetuskäyttöön mainio
 - Opiskelijoiden ei tarvitse asentaa ohjelmistoja
 - Dokumenttien jakaminen ja muokkaaminen helppoa

```

kuukaudet = data['Kk']
lämpötilat = data['Keskilämpötila (degC)']

# piirretään kuvaaja, jossa x-akselille tulee kuukaudet ja y-akselille lämpötilat
# Kuvaajan piirretään punaisella viivalla ('r-') ja datapisteet merkitään kuvaajaan
plt.plot(kuukaudet, lämpötilat, 'r--', marker='o')

# Asetetaan x-akselin ja y-akselin otsikot
plt.xlabel('Kk')
plt.ylabel('Lämpötila (°C)')

# Asetetaan kuvaajalle pääotsikko
plt.title('Keskilämpötila Helsingissä kuukausittain 2020')

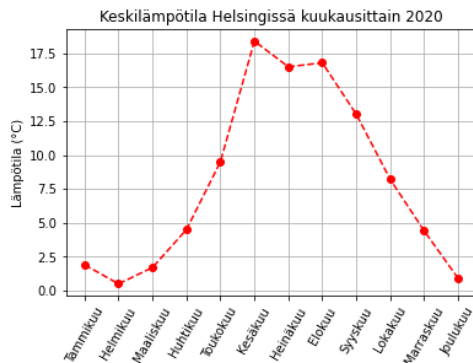
# Tehdään lista x-akselille tulevista teksteistä.
labels = ['Tammikuu', 'Helmi-kuu', 'Maaliskuu', 'Huhtikuu', 'Toukokuu', 'Kesäkuu', 'Heinäkuu', 'Elokuu', 'Syyskuu', 'Lokakuu']

# Muutetaan x-aksella olevat numerot äsken määrittämiksi teksteiksi.
# range(1,13) luo listan, jossa on numerot 1-12.
# Seuraavaksi annetaan parametrina tekstit, jotka sijoitetaan numeroiden paikalle (tallennettu 'labels'-muuttujaan)
# Lopuksi vielä käännetään tekstejä, jotta ne eivät menisi päällekkäin.
plt.xticks(range(1,13), labels, rotation=60)

# Näytetään kuvaajan ruudukko
plt.grid()

# Tämä komento poistaa ylimääräiset tulosteet ja näyttää ainoastaan kuvan, kun solu suoritetaan.
plt.show()

```



Python-ohjelmointi

- Aloittelijaystävällinen ohjelmointikieli
- Laajasti käytössä mm. tieteellisessä yhteisössä ja data-analyysissä

```
In [1]: print("Hello, World!")  
Hello, World!
```

Tutustutaan data-analyysiin notebook-ympäristössä seuraavan esimerkin avulla:

<https://mybinder.org/v2/gh/opendata-education/Python-ja-Jupyter/main?urlpath=tree/materiaali/harjoitukset/jupyter-intro.ipynb>

Kokeile itse!

- Lisää uusi tekstisolu ja lisää siihen tekstiä
- Lisää uusi koodisolu, tee jokin laskutoimitus ja suorita koodi
- Ota mallia esimerkkikoodista ja piirrä jokin kuvaaja, esimerkiksi:
 - Tampereen keskilämpötilat tiedostosta `keskilampotila2020_Tampere.csv`
 - Oulun keskilämpötilat tiedostosta `keskilampotila2020_Oulu.csv`
 - Histogrammi lumen syvyydestä Helsingissä (käytä eri saraketta datasta kuin esimerkissä)

Kotitehtävä

- Kertaa Jupyter Notebookin ja Pythonin käyttöä käymällä seuraava notebook rauhassa läpi samalla koodaten ohjeiden mukaan:
 - https://mybinder.org/v2/gh/opendata-education/Python-ja-Jupyter/main?urlpath=tree/materiaali/harjoitukset/data-analyysi_esimerkki.ipynb
- Hae itse ilmatieteenlaitoksen hakupalvelusta dataa ja piirrä siitä kuvaaja tai histogrammi
 - <https://www.ilmatieteenlaitos.fi/havaintojen-lataus>
 - Voit käyttää esim. edellisessä kohdassa olevaa notebookia