

## 1. Henkilötiedot

Numeerisen datan visualisointikirjasto; Helena Hilander, 655031,  
Bioinformaatioteknologia, 2017, 20.2.2018

## 2. Yleiskuvaus

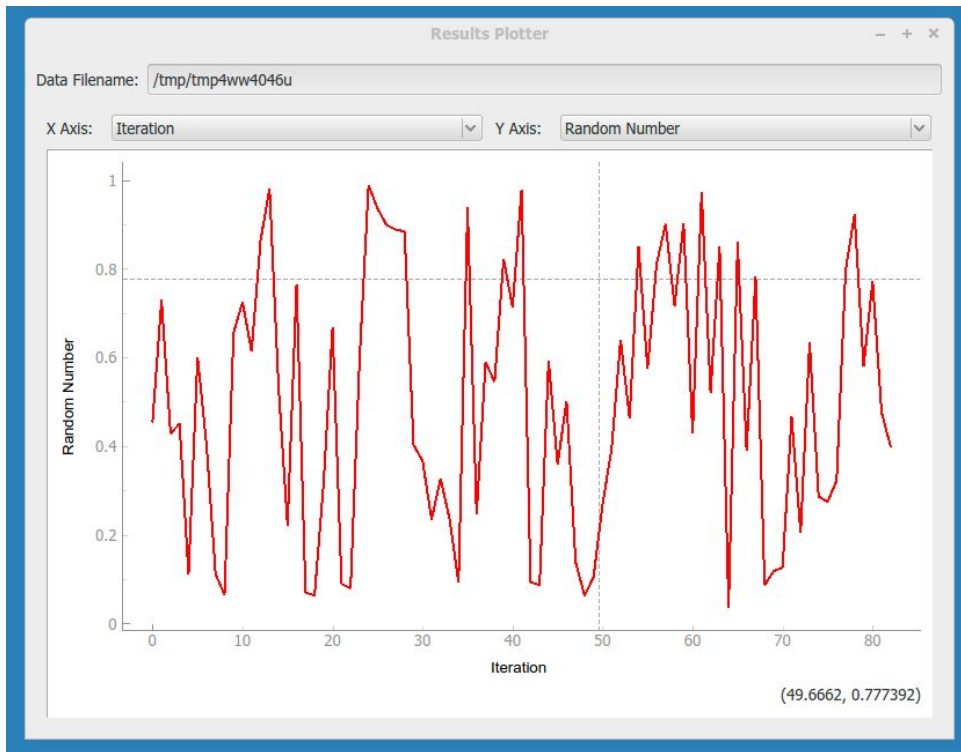
- toteutetaan ohjelmakirjasto, jolla voidaan visualisoida numeerista dataa graafisesti.
- pyrin tekemään työn vaativan version, joten perusviivadiagrammi, peruspylväsdiagrammi ja piirakkadiagrammi tulisi pystyä esittämään.
- datatiedosto luetaan käyttäjältä.
- syötetylle datalle määritellään formaatti, jota sen tulee noudattaa, jotta ohjelma toimii. (määrittely alla)
- laaditaan esimerkkisovellusdata
- **Perusviivadiagrammissa:**
  - piirretään 1-n käyrää, joilla jokaisella on oma numeerinen x-y data. Jokaisen kuvaajan omat datapisteet yhdistetään viivoilla.
  - jokaisella käyrällä on oma nimensä/selitteensä joka luetaan datasta. Eri käyrät piirretään eri väreillä
  - x- ja y-akselien nimet luetaan käyttöliittymässä käyttäjältä
  - x- ja y-akselien numeroinnin tulee olla selkeä, eikä numerot saa mennä toistensa päälle
  - käyttäjä voi lisätä gridin
- **Peruspylväsdiagrammissa:**
  - piirretään 1-n pylvästä. Jokainen pylväs edustaa yhtä x-y koordinaattiparia. X-koordinaatti sisältää datan nimen ja y-koordinaatti arvon.
  - akselien nimet ja kuvaajan otsikko luetaan käyttäjältä
- **Piirakkadiagrammissa:**
  - yksiulotteiselle datalle (vain x-koordinaatti) lasketaan jokaisen arvon suhteellinen osuus. Jokaisen arvon esiintyvyys esitetään tietyssä osuutena ympyrän alasta.
  - kukin osa ympyrästä tulee olla eri värillä

### 3. Käyttöliittymän luonnos

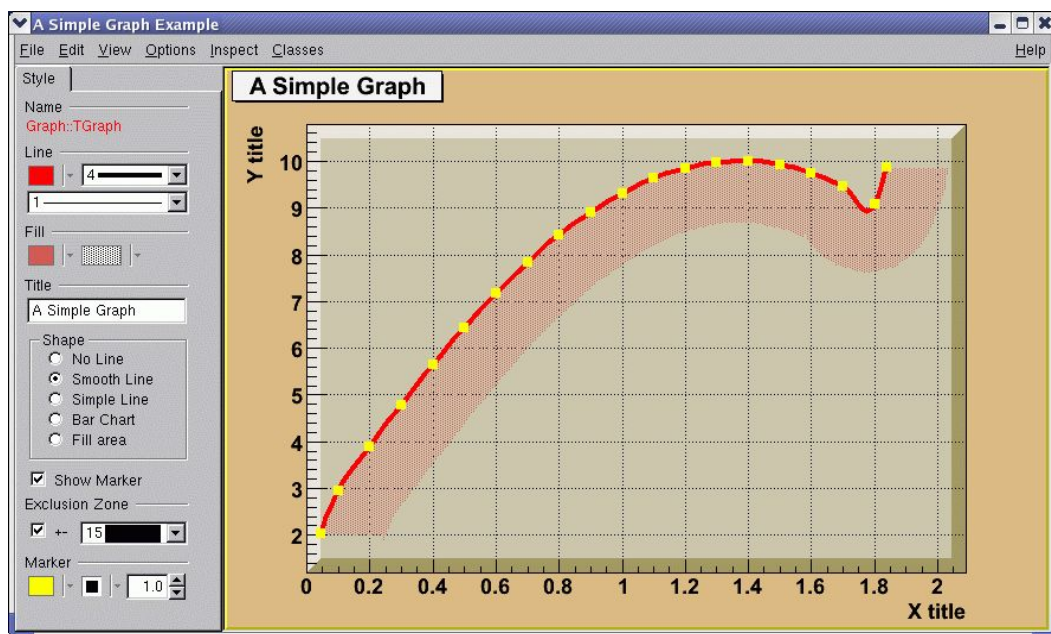
- Ohjelma kommunikoi käyttäjän kanssa graafisen käyttöliittymän avulla.
- ohjelman tuntemat komennot näkyvät käyttäjälle graafisen käyttöliittymässä nappeina ja tekstikenttinä (käyttöliittymäkomponentteina).

Tarkemmin käyttöliittymän ulkonäöstä:

- Graafisessa käyttöliittymässä tiedoston luku tapahtuu yläpalkin "open file" komennolla, joka avaa erillisen ikkunan tiedoston luvulle, Tiedoston nimi syötetään tai haetaan koneelta browsella. Tiedoston luku aktivoidaan "ok" napista.
- tiedoston luvun onnistuessa data visualisoidaan automaattisesti datatyypin mukaan oletusarvoisesti joko viiva, pylväs tai piirakkadiagrammina. Visualisointi tehdään pääikkunassa olevalle suorakulmioalueelle. Jos datatyypillä on mahdollista käyttää myös muita visuaalisia muotoja, kuvaajan tyyli voidaan vaihtaa näihin käyttöliittymän pääikkunassa olevasta "Plot type" valikosta. (esim numeerinen x-y data voidaan esittää pelkästään x-koordinaatin piirakkadiagrammina, vaikka tämä ei välttämättä analyysimielessä ole aina järkevää)
- Gridin lisääminen on linkitetty pääikkunan nappiin, jota painamalla gridin saa päälle/pois
- kuvaajan ja akselien nimeäminen tapahtuu joko pääikkunan laidasta olevista tekstikentissä tai suoraan kuvaajan akseleita tuplaklikkaamalla
- muutama esimerkkikuva, millaista ulkomuotoa ohjelmaruudulle pyrin saamaan:



lähde: <https://pymeasure.readthedocs.io/en/latest/tutorial/graphical.html> (20.2.2019)



lähde: <https://root.cern.ch/root/html/guides/users-guide/ROOTUsersGuide.html> (20.2.2019)

#### 4. Tiedostot ja tiedostoformaattit

- Käyttäjältä luetaan .csv tai .txt -tiedosto, jossa jokainen #datan\_nimi -rivillä alkava lohko kuvaa yhtä mittausta. "datan\_nimi" paikalla on kyseisen datan nimi, joka kopioidaan viivadiagrammiin käyrän nimeksi.
- Tiedoston alussa voi olla kommentteja, joita ei huomioida.
- Rivin tulee olla muotoa: x-koordinaatti tai x-koordinaatti , y-koordinaatti. Välilyönneillä ei ole merkitystä. Ohjelma tunnistaa, mikäli data on yksiulotteista, eikä kaadu pilkun puuttumiseen. Ohjelma myös tunnistaa (tai kysyy käyttäjältä) datan tyyppin. (yksiulotteinen, numeerinen kaksiulotteinen, kategorinen kaksiulotteinen)
- Jokainen tiedostosta luettu erillinen # alkuinen lohko tallennetaan omaksi omaksi Data luokan olioksi, jolla on ominaisuuksinaan mm. x-koordinaatti, y-koordinaatti, datan pituus, ja laatu.
- Mikäli yhdessä tiedostossa on useampi datasetti, jokaisen datasetin tulisi olla samaa muotoa (yksiulotteinen, numeerinen kaksiulotteinen, kategorinen kaksiulotteinen)

#### 5. Järjestelmätestaussuunnitelma

Tiedoston käsittelyyn liittyviä virhetilanteita:

- syötetään väärä tiedostotyyppi (ilmoitus käyttäjälle)
- tiedostossa ei ole yhtään # -lohkoa (ilmoitus käyttäjälle)
- liikaa koordinaatteja rivillä (ohjelman ei pitäisi tulostaa virheilmoitusta, vaan se huomioi vain 2 ensimmäistä koordinaattia, ja jatkaa toimintaa)
- jollain rivillä toinen koordinaateista puuttuu (virheilmoitus)
- sama x-koordinaatti yhdessä datassa useaan kertaan, jolloin viiva/pylväsdiagrammin teko ei onnistu. (todnäk. toteutetaan niin, että tässä tilanteessa ohjelma huomioi ainoastaan ensimmäisen kerran, kun x-koordinaatti esiintyy, ja jättää muut huomiotta, ilman, että ohjelman suoritus keskeytyy)
- samassa tiedoston eri lohkoissa erimuotoista dataa (ilmoitus käyttäjälle)

graafiseen käyttöliittymään liittyviä virhetilanteita:

- vanha kuvaaja ei saa jäädä kuvaikkunaan, kun uusi tiedosto ladataan, vaikka ohjelmaa ei suljettaisi välissä
- käyttöliittymässä tulee näkyä vain kyseiselle datatypille sopivat visualisointivaihtoehdot