

Data analytics for personalized genomics and precision medicine**Lecture2 scribing**

Lecturer: YuLI

Friday, 9 September 2022

Scriber: Kam Hei Man 1155144587

Different data types:

- Sequential data: data with sequences (1,2,3,4...) e.g. DNA sequence
- Data matrix: data expressed as matrix; can interchange or delete the whole row/column without affecting other data (e.g. table)
- Spatial data: data can be expressed as 2D/3D; spatial sequence matters so cannot interchange or delete the whole row/column that will affect the completeness of data (e.g. photo)
- Temporal data: data changed with time (e.g. stock, ECG)
- Graph/networks: data that shows the connection/link between objects (e.g. social network)
- Text: sentences
- Multi-modality data: mix of data types (e.g. video: spatial data + text + ...)
- Unknown data type: data with undefined data type

Introduction to Python programming:

- If
Python = wechat software,
Python plug-in = wechat mini program that help give us more functions (e.g. numpy, scipy, panadas)
- Python plug-in needs to load before use i.e. type: import numpy
- Examples of some python coding:
 1. `a = [1,2,3] ##store[1,2,3] in box name a`
 2. `numpy.mean(a) ##find mean value of a`
`##use plug-in program named numpy`
`##use "find mean value" function from plug-in program`

3. `print("array is", a)`
gives->array is [1,2,3] `##print everything in ()`
`##built-in python code that do not require plug-in`
`(NO!! xxx.print NO!!)`
4. `help(str)` `##` if we do not know any code's meaning, type `help()`