

Pythonで位置データを扱うハンズオン 最初のパート

2022-09-16 合同会社 長目
小川 英幸

GIS

- Geographic Information System
- 地理情報システム
- 位置に関する情報を持つデータを触る技術

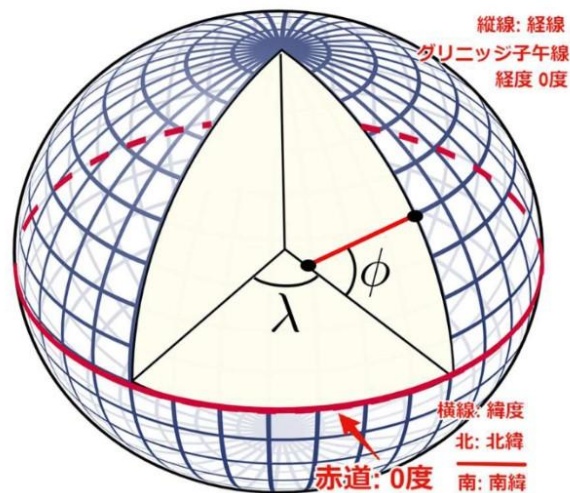
- おすすめページ
 - GISやるならやっつけ: <https://github.com/Automating-GIS-processes/site>
 - 国交省 国土地理院: <https://www.gsi.go.jp/GIS/whatisgis.html>
 - ESRI ジャパン: <https://www.esri.com/getting-started/what-is-gis/>
 - GIS実習オープン教材: <https://gis-oer.github.io/gitbook/book/>

位置データといえば住所じゃない？ => じゃない！！座標で表現する(変換をジオコーディングと呼ぶ)

- 座標: wikipedia : <https://ja.wikipedia.org/wiki/%E5%BA%A7%E6%A8%99>
 - 点の位置を指定するために与えられる数字の組(例: 緯度・経度)
- 座標系
 - 地球は楕円なので、数値で表現するにはいろいろ努力が必要。
 - 座標を表現するための、原点などの設定方法(例: 緯度経度は、赤道と本初子午線を基準として、南北を90度・東西を180度・・・)
 - いろいろある。EPSG: <https://epsg.io/>
 - これを揃えて作業しないと、とんでもないことになることもあるので、気を付ける
 - 日本だけでもたくさんあることがまとまっているページ:
<https://lemulus.me/column/epsg-list-gis>
 - 緯度経度、小数ではなく度・分・秒表現とかもあるので注意！

座標

- 緯度経度
 - 緯度
 - 南北を表現（90度まで）
 - 赤道を0度
 - 経度
 - 東西を表現（180度まで）
 - 本初子午線を0度: グリニッジ天文台から102. 478m東
- 小数で表現される場合もあるが度・分・秒で表現されることも多く、分秒を60で割る必要
- 緯度経度を数値で表現すると位置が特定できる
- 一方で、表現が複数存在し・・・



Peter Mercator, Public domain, ウィキ
メディア・コモンズ経由で

Pythonのパッケージエコシステム

- shapely: 座標を扱う
 - 点・線・面
- geopandas: まとまった座標を扱う(ベクターデータ)
- xarray: 画像内にデータが格納されているとかいうやつをまとめて扱える(ラスタデータ)
- これらを有名な奴たちが支えているので、興味のある方は調べてみてください
- folium / pydeck / plotly
 - 位置データの可視化
 - mapboxのトークンとかいらないのでfoliumが便利

ファイル形式

2つのデータ形式

- ラスタデータ
 - 格子状にデータが格納されている
 - ファイルフォーマット: GeoTIFFなど
 - ライブラリ: rasterio, **xarray**
- ベクタデータ
 - ポイントやラインストリング、ポリゴンから作られる座標データ
 - ファイルフォーマット: Shape、GeoJSON、KMLなど
 - Shapeは3-5個くらいのファイルから構成される
 - ライブラリ: **geopandas**

つづきはのーとぶっくで！

- 参考資料
 - 小川の去年のPyConJPでの発表資料
 - <https://github.com/mazarimono/pyconjp2021>
 - **PLATEAU Hands-on 02GISでの3Dビジュアライゼーション**
 - **株式会社アナザーブレイン 代表取締役 久田 智之**
 - <https://docs.google.com/presentation/d/1uHgjsxrsE25LR42v7219RGT7sAM-blpo-F46dP0sRVM/edit?usp=sharing>