**記者ゼミIT講座 2020/3/14 13:00-17:00**

１．ねらい

**地理情報システム（GIS）の**報道利用の具体例を紹介するとともに、**QGIS**などを使った**地理空間情報**の分析の基本やデータ加工の考え方を紹介する。

国土地理院：<https://www.gsi.go.jp/GIS/whatisgis.html>

GIS実習オープン教材：<https://gis-oer.github.io/gitbook/book/>

２．**GIS**の報道利用の具体例(13:00-14:00)

**多摩川沿い なぜ“浸水エリア”に新築が… 徹底分析しました (NHK NEWS WEB NewsUp2019/12/3)**

<https://www3.nhk.or.jp/news/html/20191203/k10012200031000.html>

NHK 報道局 社会部 今村　清人 記者

**南海トラフ地震 津波浸水想定エリアで起きていること**

<https://www3.nhk.or.jp/news/html/20200302/k10012309571000.html>

NHK 報道局 社会部 齋藤　恵二郎 記者

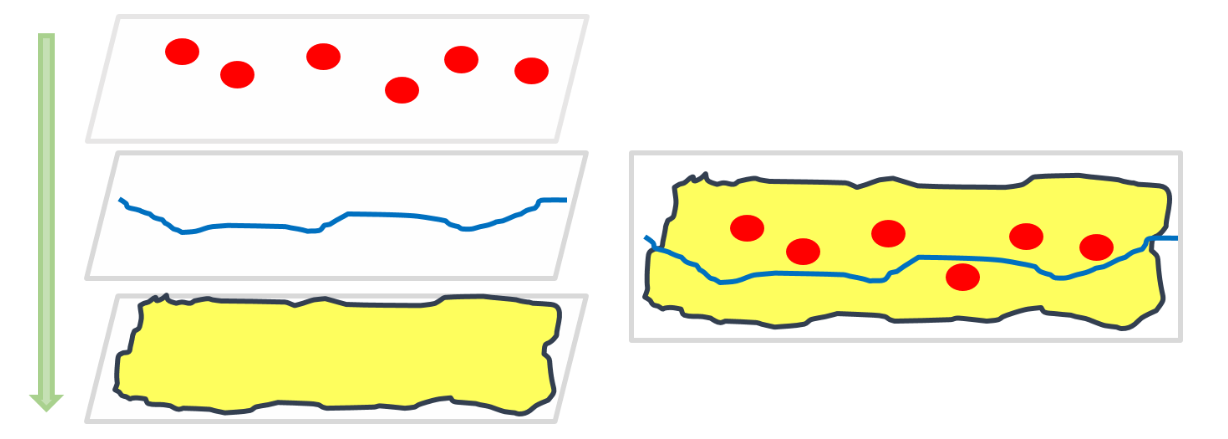
３．**GIS**の基本と利点(14:00-15:00)

* **地理空間情報**：空間上の特定の地点又は区域の位置を示す情報（位置情報）とそれに関連付けられた様々な事象に関する情報、もしくは位置情報のみからなる情報をいう。

（例）土地利用図、地質図、ハザードマップ等の主題図、都市計画図、地形図、地名情報、統計情報、空中写真、衛星画像等

* **地理情報システム（GIS：Geographic Information System）**: コンピューターを用いて、地理空間情報（地理空間データ）を、可視化、作成、編集、検索、分析するシステム。
* **地物(Feature)**：道路、河川、建物、境界、線路など、**地球上にあるすべてのものを示す概念。**
* GISではデータ**（レイヤ）を重ね合わせて**地物を表現する。

<https://gis-oer.github.io/gitbook/book/materials/00/00.html>



* 地理情報システムの**活用例**

1. 最短経路や特定の領域の算出
2. 建物等の分布密度の可視化
3. 古地図の位置あわせ
4. 環境変化の可視化と面積計算
5. 標高データから地形情報を算出、可視化

<https://gis-oer.github.io/gitbook/book/materials/00/00.html#%E5%9C%B0%E7%90%86%E6%83%85%E5%A0%B1%E3%82%B7%E3%82%B9%E3%83%86%E3%83%A0%E3%81%A8%E3%81%AF>

４．**QGIS**による地理空間情報分析の基本(15:00-16:00)

**１．QGISとは**

* 無償で利用できる**オープンソース**のGIS。
* Windows, Mac, Linux, Unix, Androidなどの環境で動作。さまざまなプラグインや、GRASSやPostGISなど他のオープンソースGISと連携して使用することで、**多種類の分析が可能**。
* GISの基本操作に必要な機能を網羅しているだけでなく、Web上に多数のマニュアルがアップロードされているため、**学習がしやすい。**
* データ形式：ベクターとラスタ

ベクター データとは（esriジャパン）

<https://www.esrij.com/gis-guide/gis-datamodel/vector-data/>

ラスタ　データとは（esriジャパン）

<https://www.esrij.com/gis-guide/gis-datamodel/raster-data/>

**２．実習**

多摩川水系の浸水想定区域データ（国土数値情報）と国勢調査の５００ｍメッシュデータ（e-Stat）を重ね合わせることで、浸水想定区域の浸水深ごとの人口推移を分析する。

1. 多摩川水系の浸水想定区域データ（５段階浸水深）

**→地図（shp）の追加**

（データ１：Data/Result/Shinsui/多摩川水系の浸水想定区域データ.shp）

1. 国勢調査の５００ｍメッシュデータ（1995-2015の総人口）

**→プロパティの使い方**

（データ２：Data/Result/Pop/多摩川水系を包含する国勢調査のメッシュデータ.shp）

1. １と２を交差（Intersect）して面積按分することで、浸水深ごとの人口データを取得

**→面積計算など（フィールド計算機）**

[NA-GIS104]フィールド計算機を使う

<https://na-mii.com/gis/na_gis104/>

（データ３：Data/Result/Before\_Anbun/交差データ　面積按分計算前.shp）

（データ４：Data/Result/After\_Anbun/交差データ　面先按分計算後.shp）

1. 浸水想定区域の人口推移を地図上に可視化して傾向を分析

**→QGIS操作の基本**

1. 地理院タイル地図を重ねることで、人口増加地域の土地履歴を分析

**→タイル地図のインポート**

朝日航洋：地図タイルの活用について

<https://www.aeroasahi.co.jp/qgis/post/2019/10/qgis_tile/>

（データ５：Data/Result/Image/武蔵小杉と二子玉川.png）

1. 浸水深ごとの人口推移をグラフに可視化して傾向を分析

**→CSVへのエクスポート**

QGISでラインをポイントに変換してｘｙ座標を取り出す

<https://qiita.com/ishiijunpei/items/a90385cca43ddcb91428>

（データ6：Data/Result/Tamagawa\_CSV/エクセルデータ.xlsx .csv）

地図で見る統計(統計GIS)e-Stat：<https://www.e-stat.go.jp/gis>

国土数値情報：<http://nlftp.mlit.go.jp/ksj/>

地理院タイル一覧（国土地理院）：<https://maps.gsi.go.jp/development/ichiran.html>

５．分析のための**データ整形**の方法とツール(16:00-17:00)

１．国勢調査の５００ｍメッシュデータの整形（メッシュコード：５３３９）

（Row Data）

・１９９５~２０１５の4次メッシュ（５００ｍメッシュ）データ

（データ7：Data/Raw/Tokyo\_Pop/テキストデータ.txt）

（データ8：Data/Raw/Kanagawa\_Pop/テキストデータ.txt）

　１９９５~２０１０：男女別人口総数及び世帯総数

　２０１５：人口等基本集計に関する事項

<https://www.e-stat.go.jp/gis/statmap-search?page=1&type=1&toukeiCode=00200521>

・4次メッシュ（５００ｍメッシュ）境界データ

<https://www.e-stat.go.jp/gis/statmap-search?page=1&type=2&aggregateUnitForBoundary=H>

* 地域メッシュ統計について（総務省統計局）：<https://www.stat.go.jp/data/mesh/m_tuite.html>

（データ9：Data/Raw/Tokyo\_Mesh/境界データ.shp）

（データ10：Data/Raw/Kanagawa\_Mesh/境界データ.shp）

1. 神奈川県と東京都のRow Dataをダウンロード

**★定義書も一緒にダウンロードすること（重要！）**

・国勢調査５００ｍ：テキストデータ**（.txt）**

・境界データ５００ｍ**(.shp)**

・メッシュごとのファイル数：１５

神奈川：４×（境界１＋国勢（各年）４）東京：１１×（境界１＋国勢（各年）４）

→**６０ファイル**

1. 表(Dataframe)を作成する

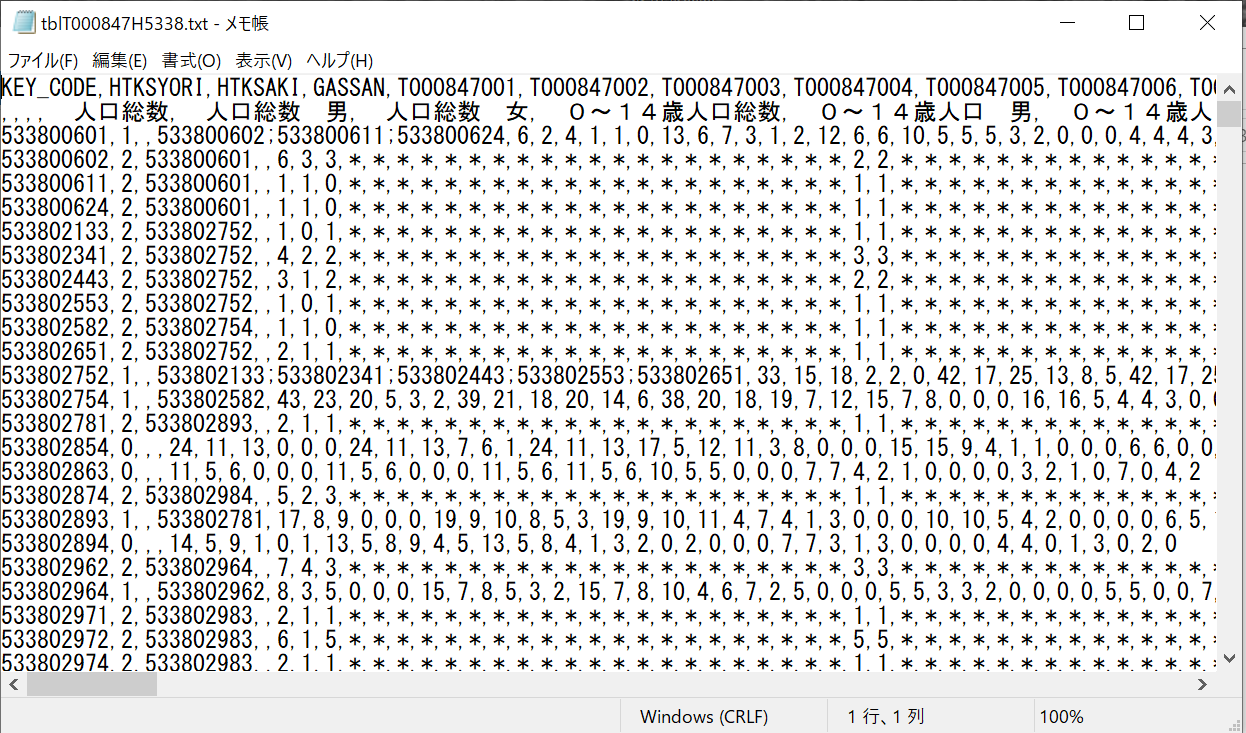
・１行目（code）と２行目（日本語）→１行目をヘッダーにする

・データ型を揃える→KEY CODEなどはテキスト(Text)に、人口は整数(Integer)に変換

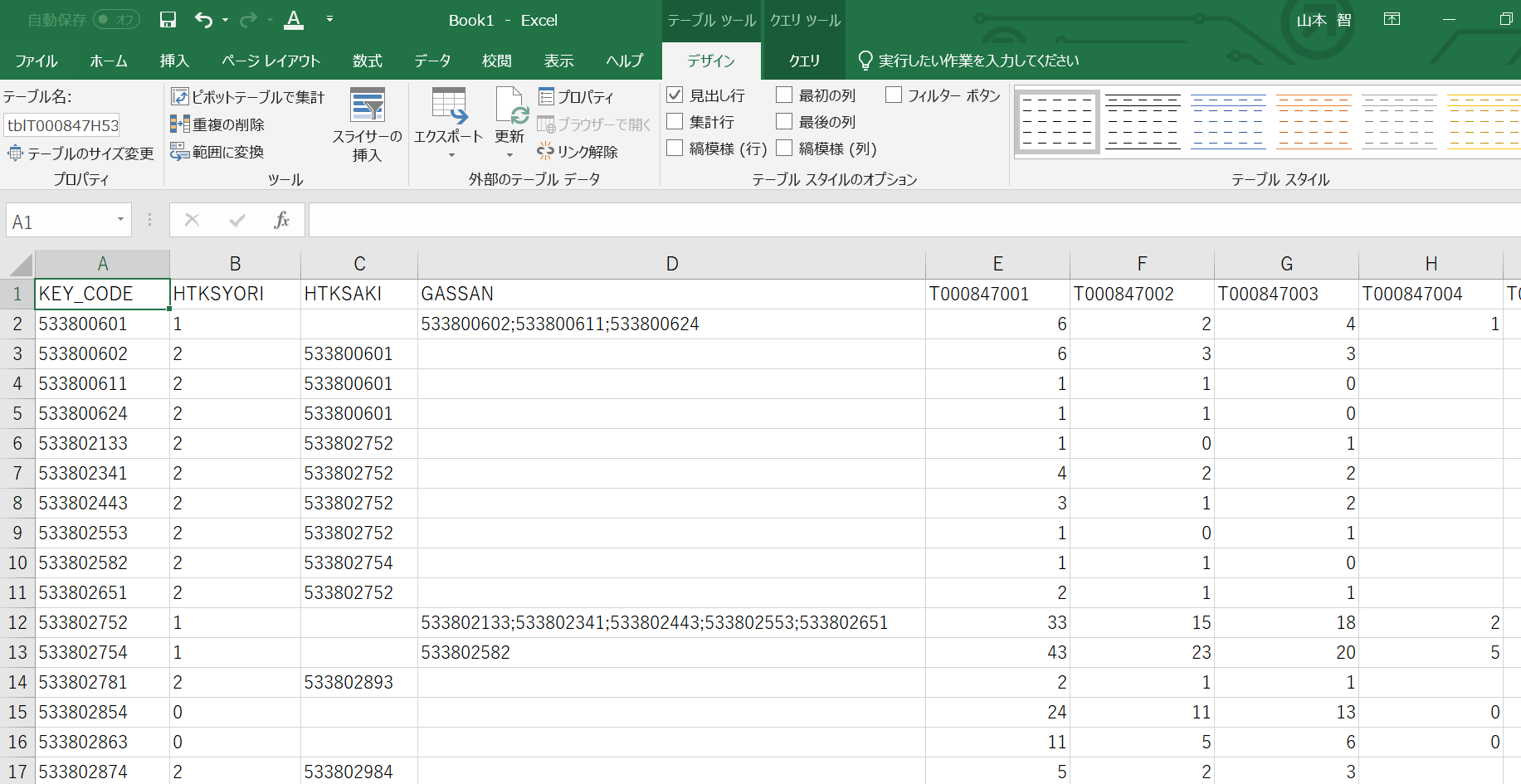
・境界データはシェープファイルなので、QGISに読み込んでCSVファイルでエクスポート

**→テキスト（CSV）のエクセルへの読み込み**

（テキストデータ）

****

（エクセルデータ）

****

1. 境界データに対して各年の国勢調査データをKEY CODEをKeyにして結合

**→エクセルの場合はVLookup関数を使う**

1. 総人口のカラム（列）だけ抽出

・"T000752001","T000386001","T000387001","T000609001","T000847001"

・N/A（No Data）の部分を０（ゼロ）に置換

1. 結合したデータの測地系を平面直角座標系に変換する

**→QGISに読み込んで変換**

・多摩川の場合は、Japan Plane Rectangular CS IX（EPSG:2451）

1. 東京と神奈川のメッシュごとに以上の手順を**１５回**繰り返す。

（データ11：Data/Working/Kanagawa\_MeshPop/各IDメッシュ人口データ.shp）

（データ12：Data/Working/Tokyo\_MeshPop/各IDメッシュ人口データ.shp）

1. メッシュコード：５３３９を採用（多摩川浸水想定区域を包含）

日本の測地系（国土地理院）：<https://www.gsi.go.jp/sokuchikijun/datum-main.html>

世界測地系（国土地理院）：<https://www.gsi.go.jp/KIDS/KIDS13.html>

JGD座標系とEPSGの一覧表：<http://blog.godo-tys.jp/2012/11/21/999/>

2．多摩川水系の浸水想定区域データ（５段階浸水深）の整形

（Row Data）

・浸水想定区域データ（都道府県ごと）（国土数値情報）

<http://nlftp.mlit.go.jp/ksj/>

1. 神奈川県と東京都のRow Dataをダウンロード

**定義書も一緒にダウンロードすること（重要！）**

・シェープファイル**(.shp)**

・属性ファイルフォルダの**テキストデータ**に"対象となる洪水予報河川"のデータが格納

　　ファイル数：神奈川（３１）東京（１１）

・**各テキスト名に番号が付与され、シェープファイルの属性テーブルに記述（A31\_006）**

（データ11：Data/Raw/Kanagawa\_Tamagawa/多摩川浸水（東京）.shp）

（データ12：Data/Raw/Tokyo\_Tamagawa/多摩川浸水（神奈川）.shp）

1. 属性ファイル内のテキストデータを**１つずつ開き**、“多摩川”の文字を検索して、番号を抽出
2. シェープファイルの属性テーブルで“多摩川”の文字がある番号のデータだけ抽出
3. 浸水深データを整形**（A31\_001）**

・地方整備局は７段階

・自治体は５段階

・**２１**→１１ ・**２２**→１２ ・**２３**→１３ ・**２４**→１４ ・**２５、２６．２７**→１５

1. 結合したデータの測地系を平面直角座標系に変換する

**→QGISに読み込んで変換**

・多摩川の場合は、Japan Plane Rectangular CS IX（EPSG:2451）

1. 東京都と神奈川県の浸水想定区域データを結合

（データ13：Data/Working/Tokyo\_Tamagawa/多摩川浸水（東京）.shp）

（データ14：Data/Working/Kanagawa\_Tamagawa/多摩川浸水（神奈川）.shp）

3．データ整形

* 1. ユニークな分析をするためにはデータの整形(Data cleansing)が必要

（データ分析の８０％は整形作業に費やされる）

* 1. Raw Dataの定義書をよく読むことが必要
  2. 整形のロジックを考えることが必要
  3. 大量のデータ洗浄はプログラムでの自動化が必要（R　Python）

Tamagawa\_Notebookファイルに、Pythonを使ったデータ整形プログラム(Jupyter notebook)が入っています。