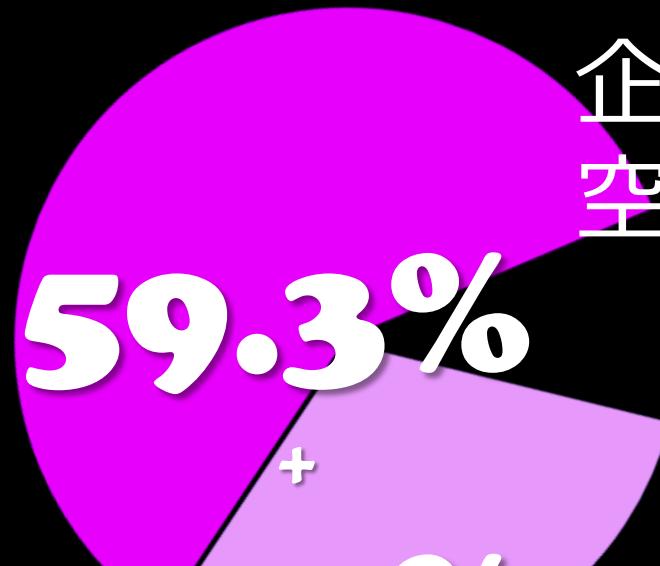


Spatial Data Scienceのはじめの一歩



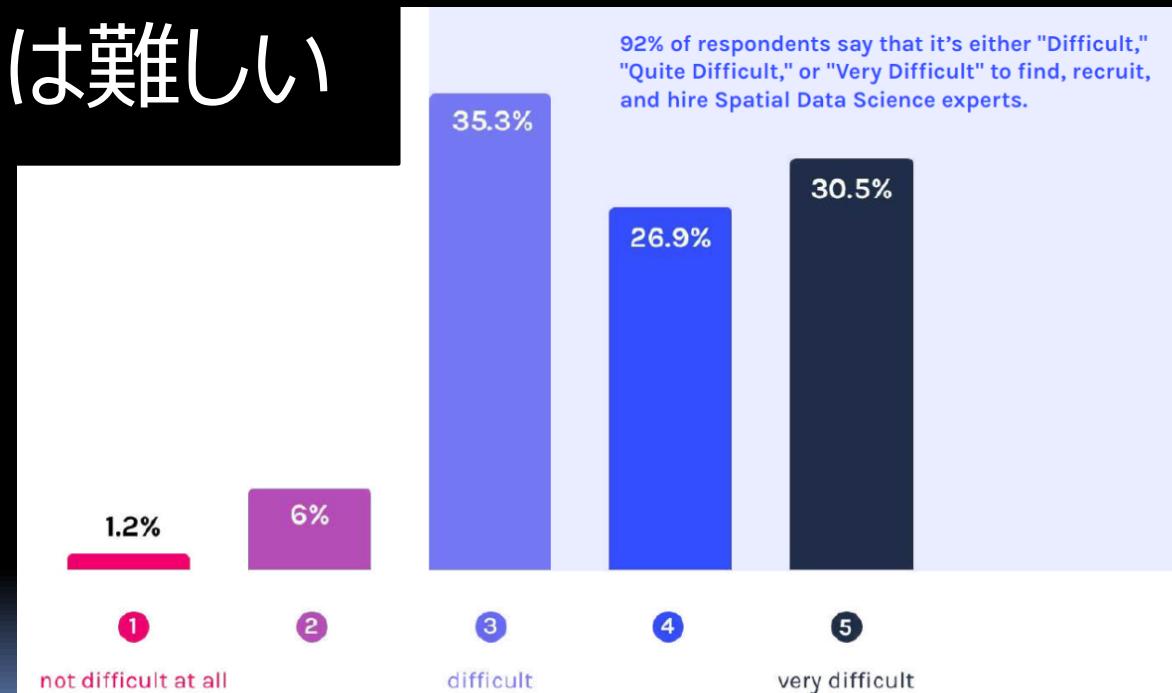
東京農工大学 大学院
グローバルイノベーション研究院
教授 赤坂 宗光

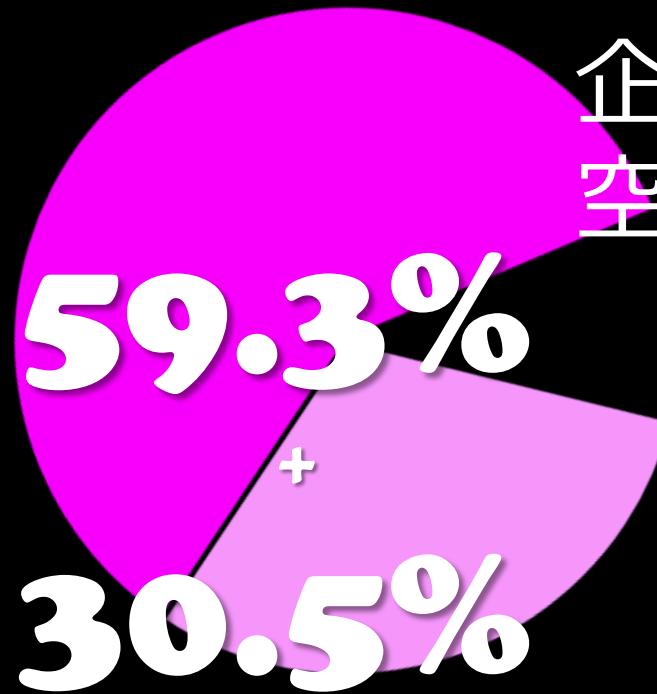


企業、公共のいずれのセクターでも約60%の組織が、
空間データサイエンスへの投資を高い確度で増やす
-約30%も増やす可能性がある

(State of Spatial data science 2023)

92%の組織が空間データサイエンスの専門家を
探し・雇用するのは難しい





企業、公共のいずれのセクターでも約60%の組織が、
空間データサイエンスへの投資を高い確度で増やす
-約30%も増やす可能性がある

(State of Spatial data science 2023)

活用が期待される分野の例

小売業：出店計画、売上予測、需要予測、競合解析

通信業：基地局の最適化、営業の効率化

配送業：配送ルートの最適化

広告業：ジオ/エリ亞ターゲティング広告

Spatial Data Science?

Geographic Data Science とも
Data Scienceの領域のひとつ

Data Science: データから意味のある情報を抽出する科学

Spatial Data Science: 空間データから
意味のある情報を抽出する科学?



Photo: <https://spatial.uchicago.edu/directory/luc-anselin-phd>

The "true" spatial data science treats location, distance, and spatial interaction as a core aspect of data and employs specialized methods and software to store, retrieve, explore, analyze, visualize and learn from such data.

-- Prof. Luc Anselin

「真の」空間データサイエンスは、データの位置、距離、および空間的な相互作用を中心的な側面として扱い、そのようなデータを格納し、取得し、探索し、分析し、視覚化し、学びを得るために、専門的な手法とソフトウェアを用いる。

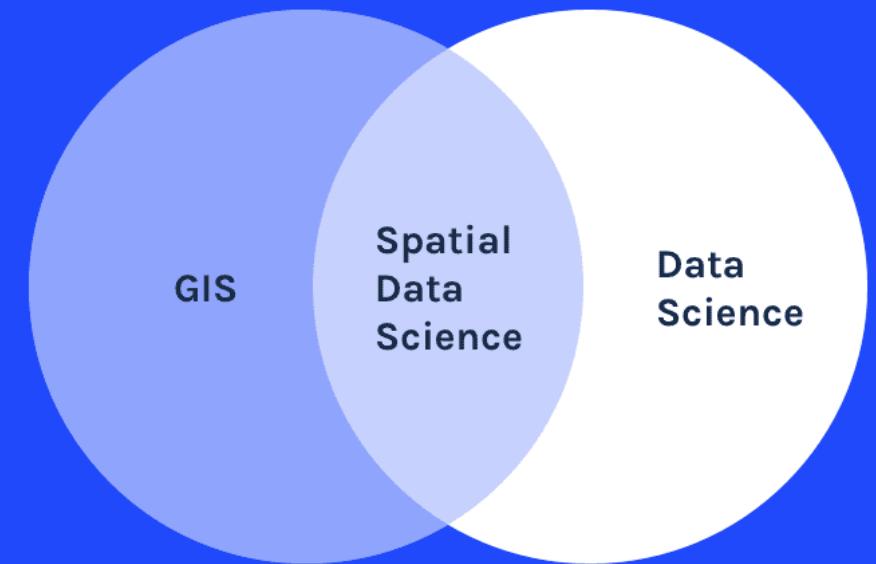


<- Prof. Anselinの空間データサイエンスの講義

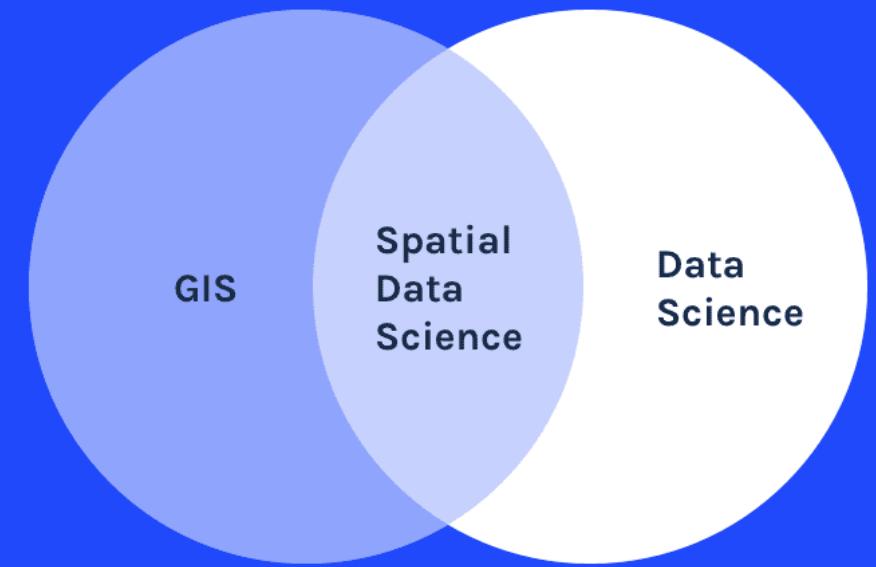
「どこで」関心のある事象が起きているか

+

関心のある事象が「なぜそこ」で起こるか

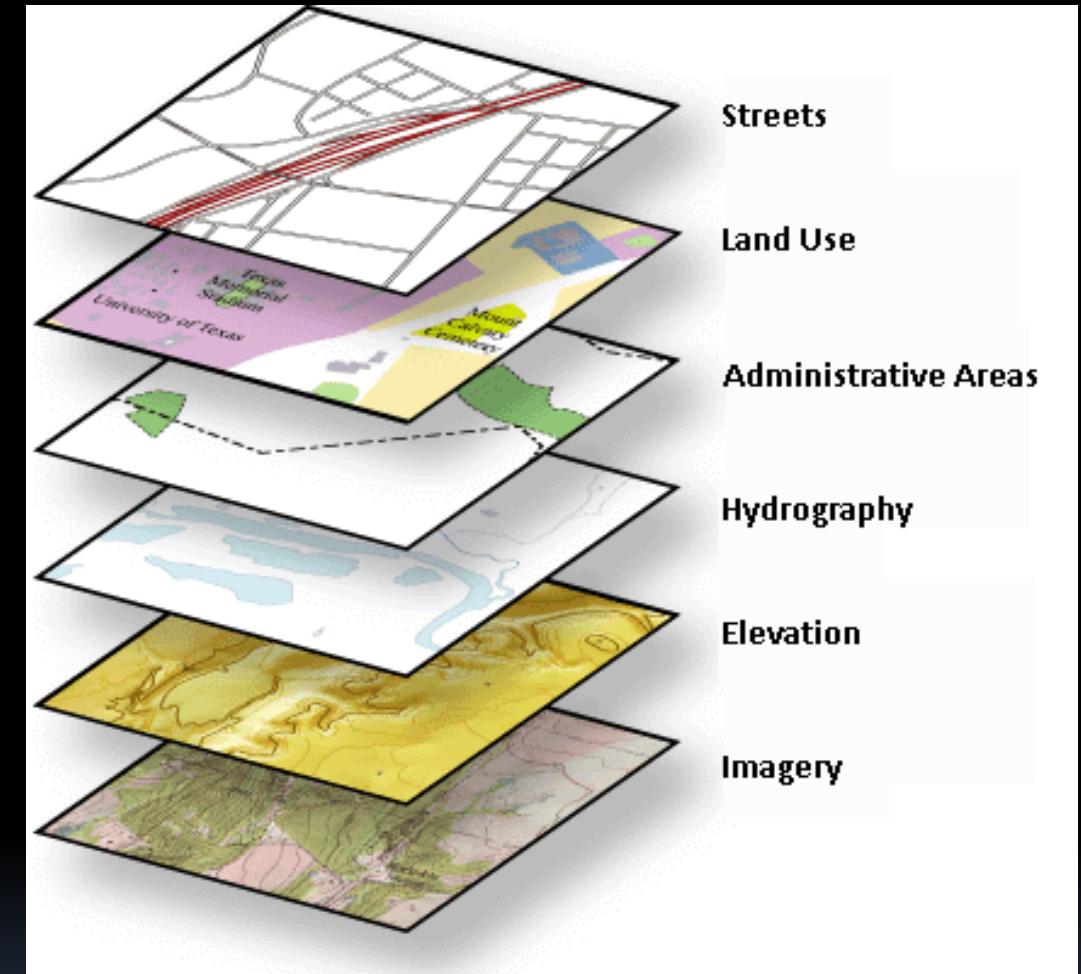


- ✓ データサイエンティストの**1/3**のみが
空間データを扱っている(State of Spatial data
science 2023)
- ✓ (もったいない)空間データの扱われ方
 - 位置情報は数ある属性の一つとして扱
われ、深堀りされていない
 - 可視化のみに利用



空間データの特徴

- 位置（緯度経度）
- 面積・体積
- 距離（時間距離）
- 経路
- 重なり度合い、交差、内包



gis-rs-2012.blogspot.com/2012/11/key-aspects-of-gis.html

Types of data

• 点データ

- ✓ 定点データ: 観測地点でのモニタリング結果（測定値・イベント発生の有無）
地震の震度、アメダス、病院毎のコロナの罹患者数、カード等の決済情報
- ✓ 一時な点データ：スマホのGPS情報など、移動する点の情報
人流データ、POI (point of Interest) データ、カード等の決済情報

• 面データ

- ✓ 明確な境界を有する面データ（境界はグリッド状/不定形）：
土地被覆、人口動態データ

• ネットワークデータ

- ✓ 点(ノード)とそれらをつなぐリンク/エッジからなるデータ
道路網、ソーシャルネットワークなど、

データの空間構造を考慮して分析するご利益

高精度で偏りの少ない予測・推定

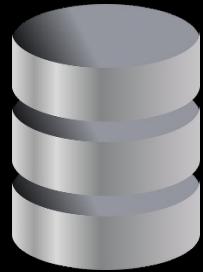


新たな次元での洞察が可能となり意思決定の精度が向上

漏出効果(sillover)を含む近隣との相互作用の程度や範囲の定量化

新しい観点での分析が可能な実施

小売業

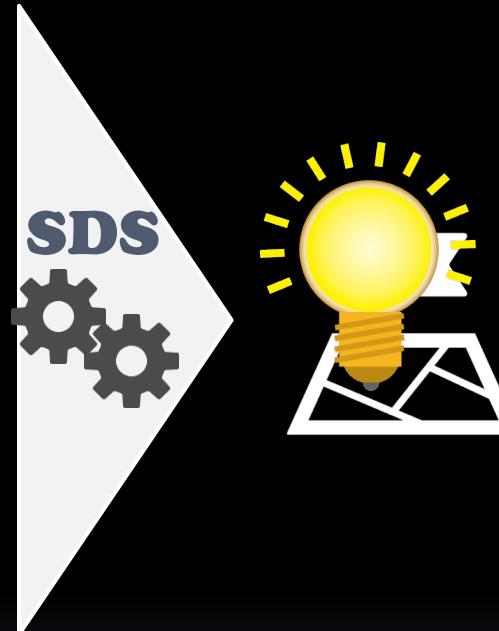


人口密度

交通パターン

競合他社の店舗との距離

気象条件

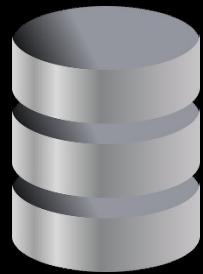


出店計画

需要予測

競合解析

不動産業



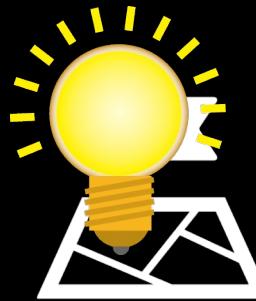
地形

土地利用規制

日用品店からの距離

災害ハザードマップ

人口動態

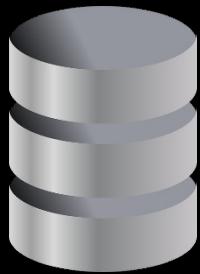


需要予測

建設適地の選定

投資対象のポテンシャル評価

物流・運輸業



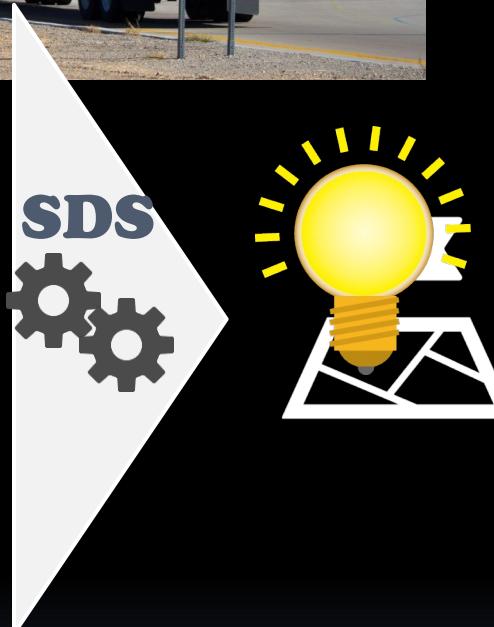
交通網

交通量・速度

地形

燃費

走行位置

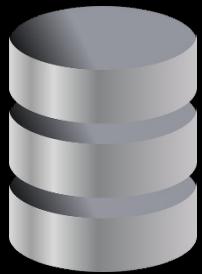


移動時間と使用燃料の最小化、
配達量の最大化する

渋滞予測・回避

最適な配送経路の把握

農業



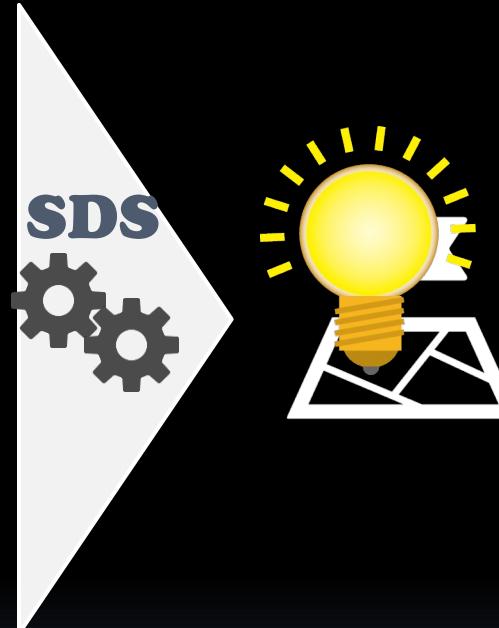
土壤の質

地形

気象パターン

作物の生育状況

病害虫の発生状況



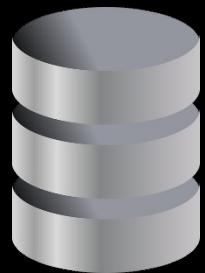
収益を最大化する

施肥パターン

農薬散布パターン

収穫時期

環境管理



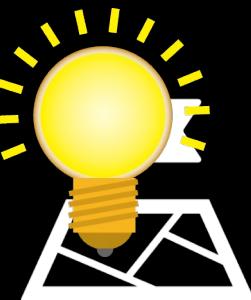
過去の違法伐採の分布

地形

交通網

(違法)市場までの距離

地域の平均収入

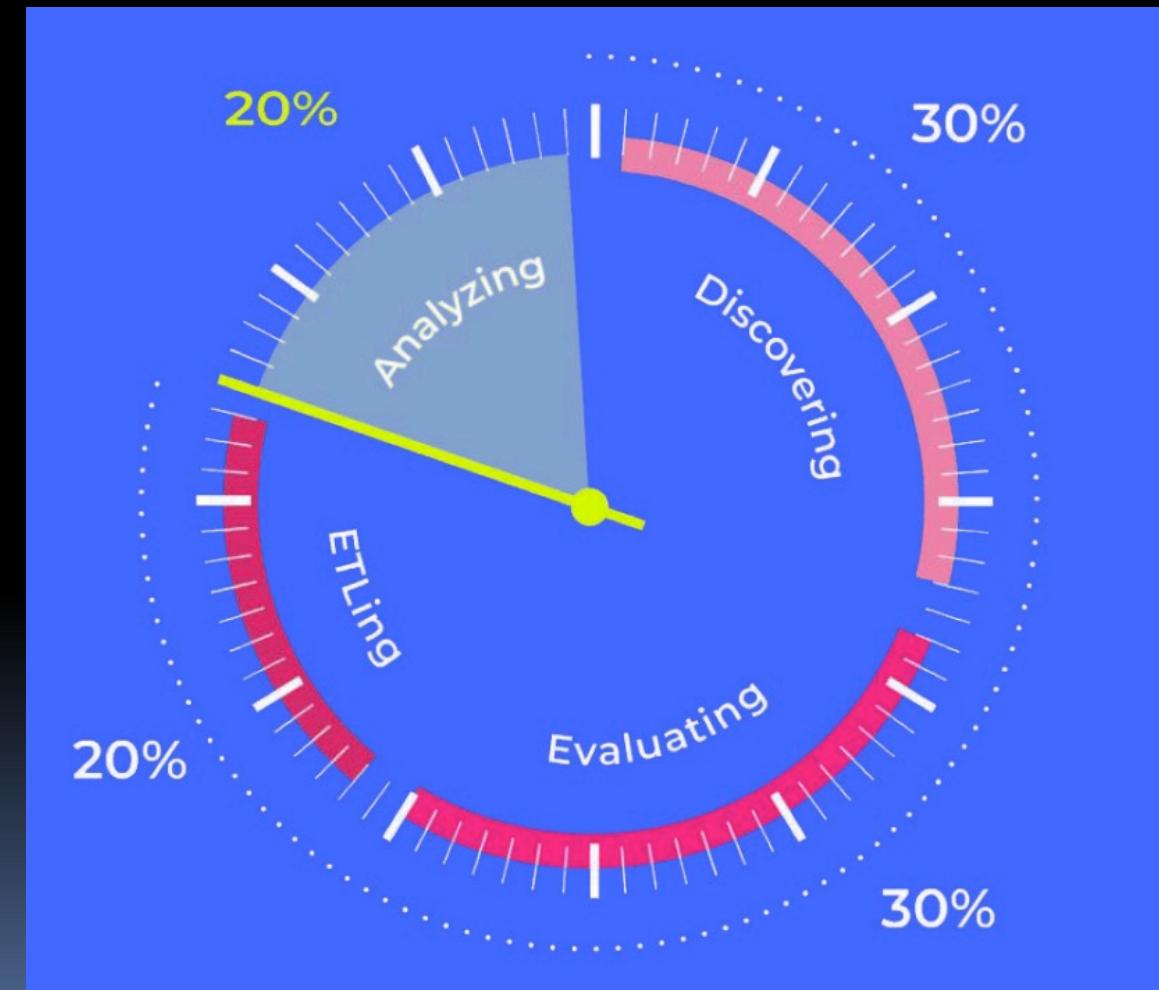


違法伐採の取締の
重点区域の特定

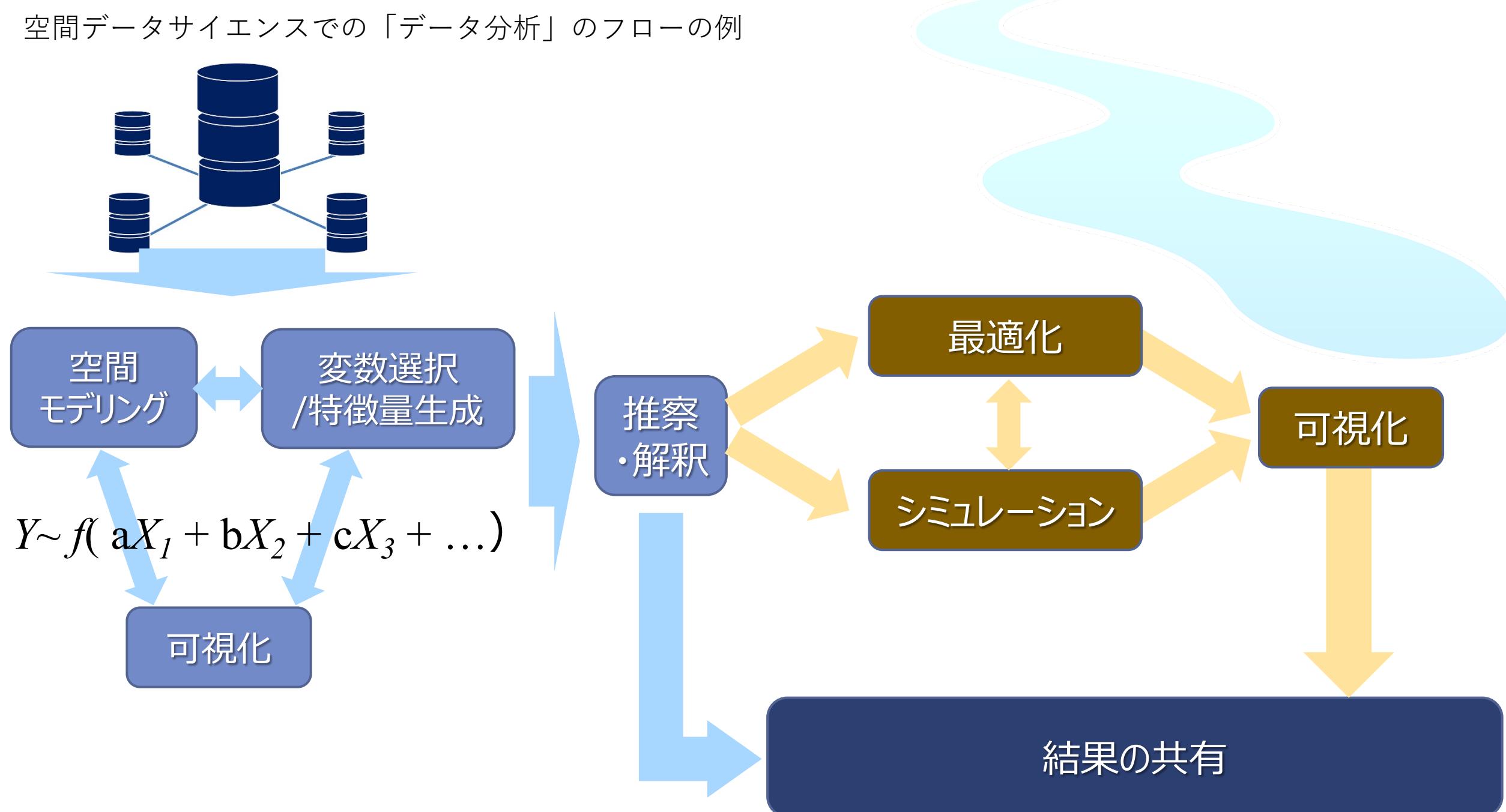
空間データサイエンスで扱うこと

- データ操作 (manipulation)
- データ統合 (integration)
- データ探索 (exploration)
- 可視化(Visualization)
- データ分析 (Analysis)

20%



空間データサイエンスでの「データ分析」のフローの例



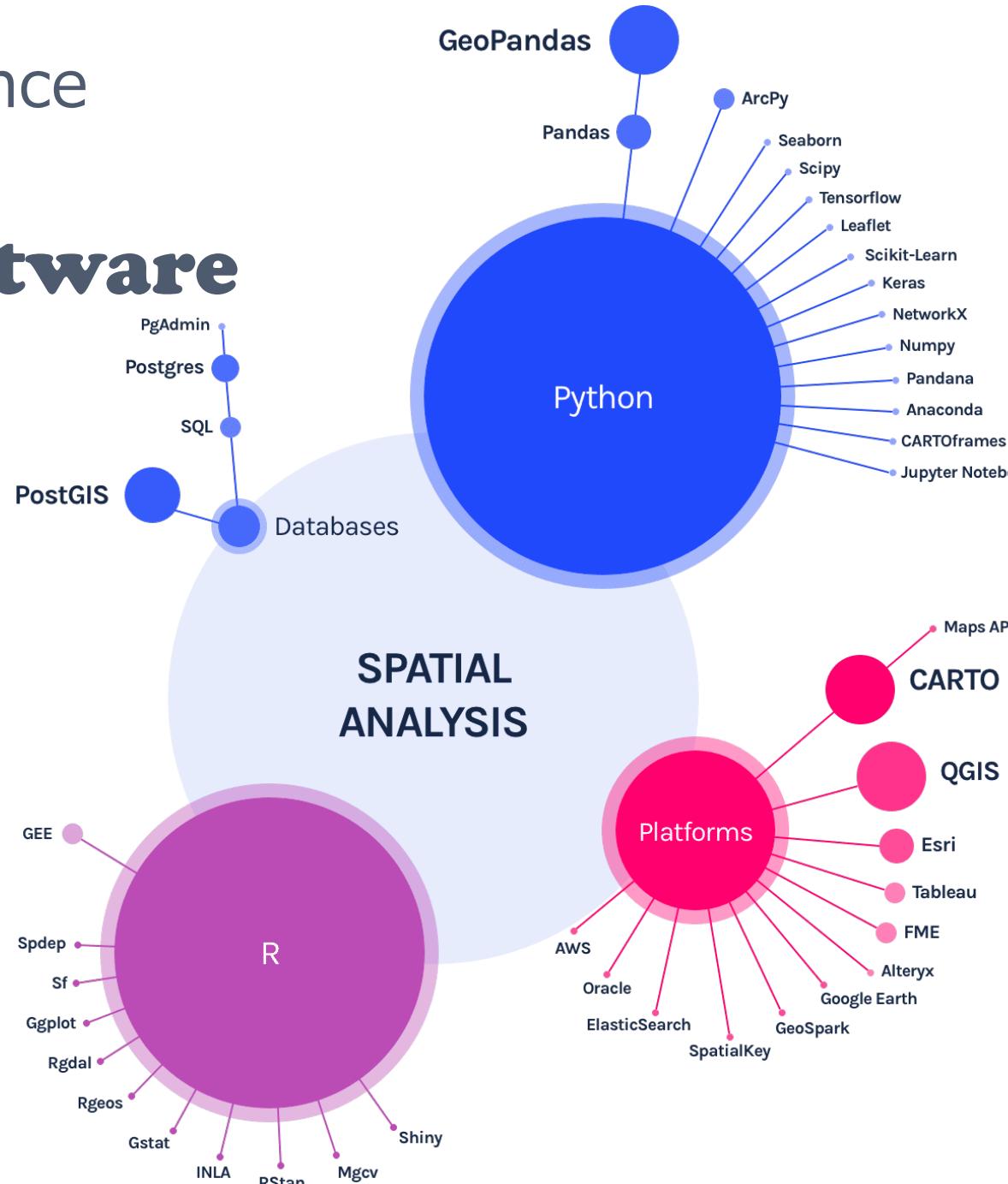
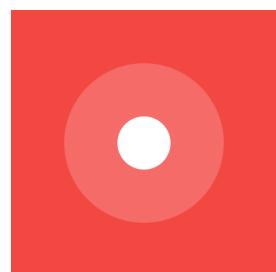
Toolbox/Skillset for Spatial data science



Photo:
spatial.uchicago.edu
/directory/luc-anselin-phd

No one size fit all software

- *Luc Anselin*



空間データを扱う際の落とし穴

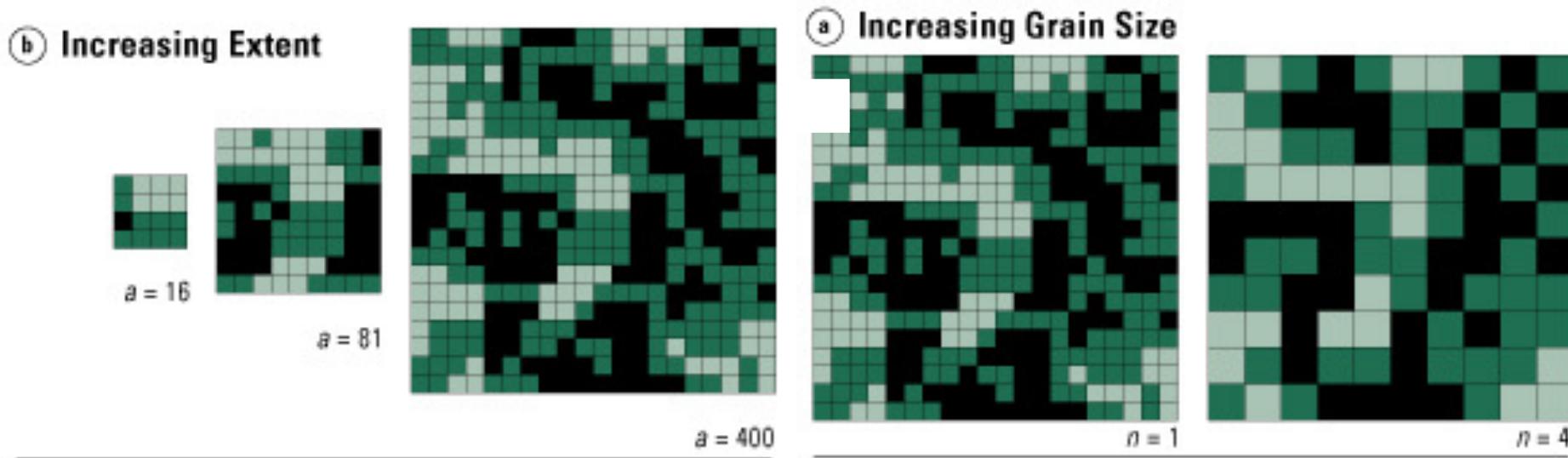


- 「注目している現象」と「データ」のスケールの不一致
- 現象・現象を規定する要因のスケール依存性

そもそも「スケール」とは？

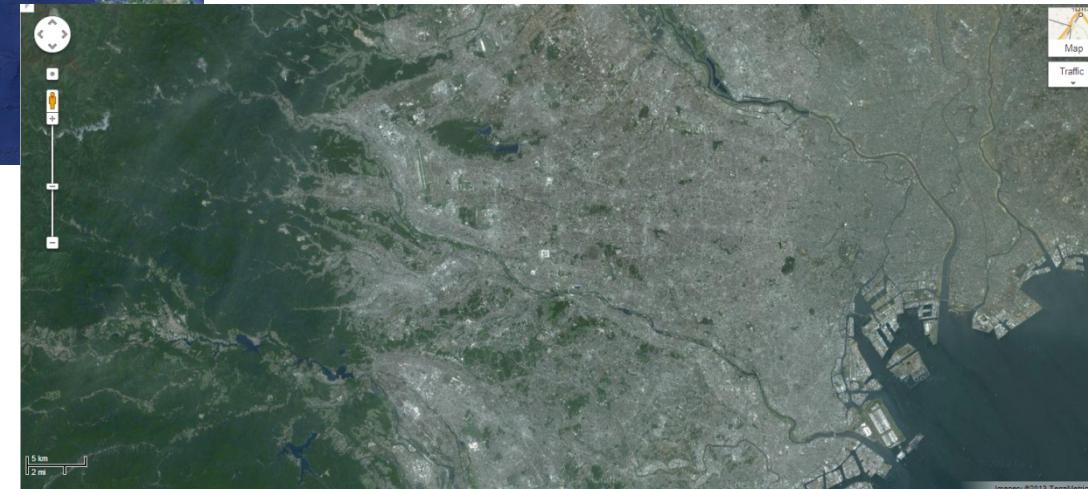
空間情報を扱う際の「スケール」

- **範囲(extent)**: 空間の**広がり**
- **解像度(grain/resolution)**: 空間の**きめ細かさ**

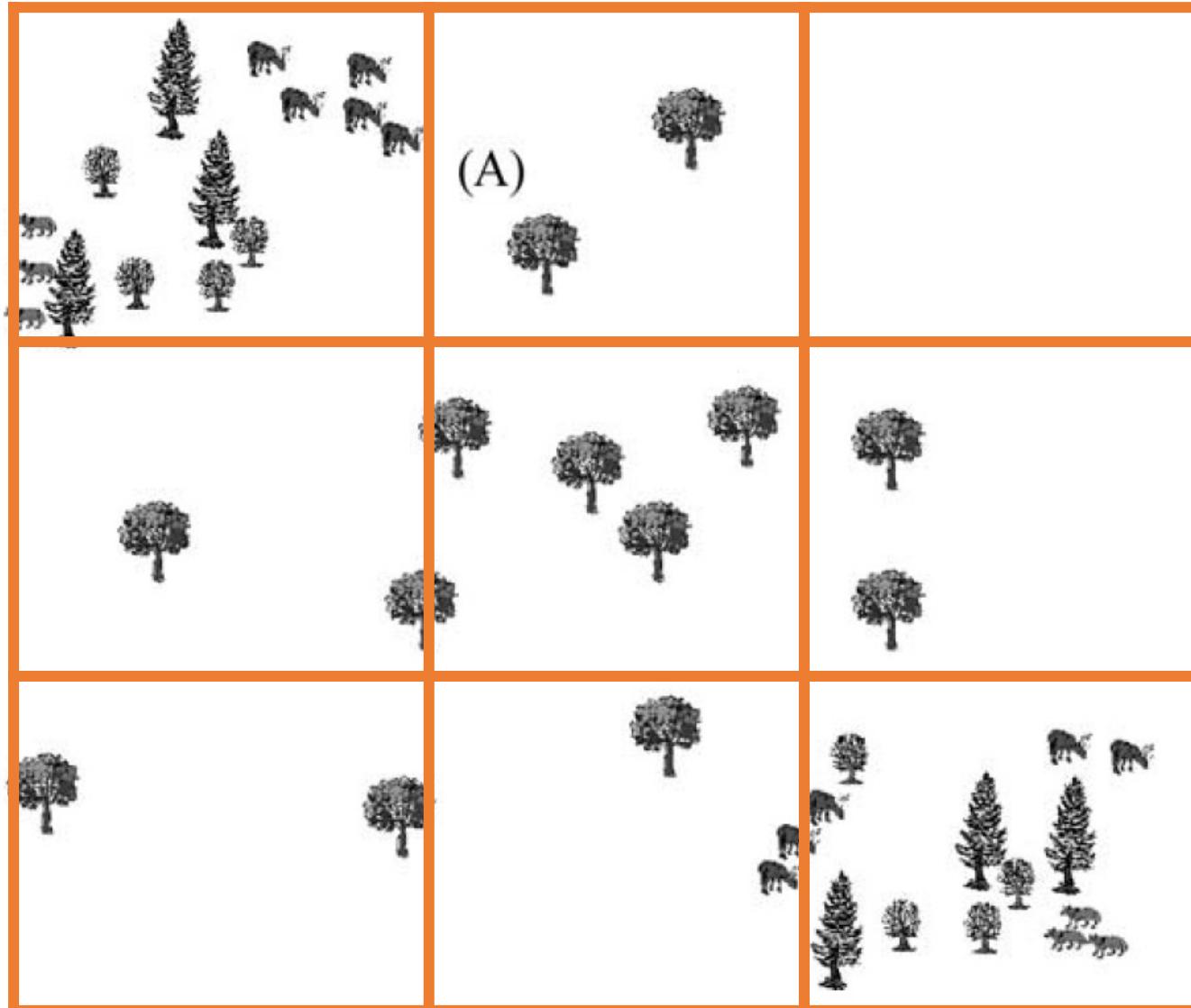


空間範囲(spatial extent)の違いは何をもたらす？

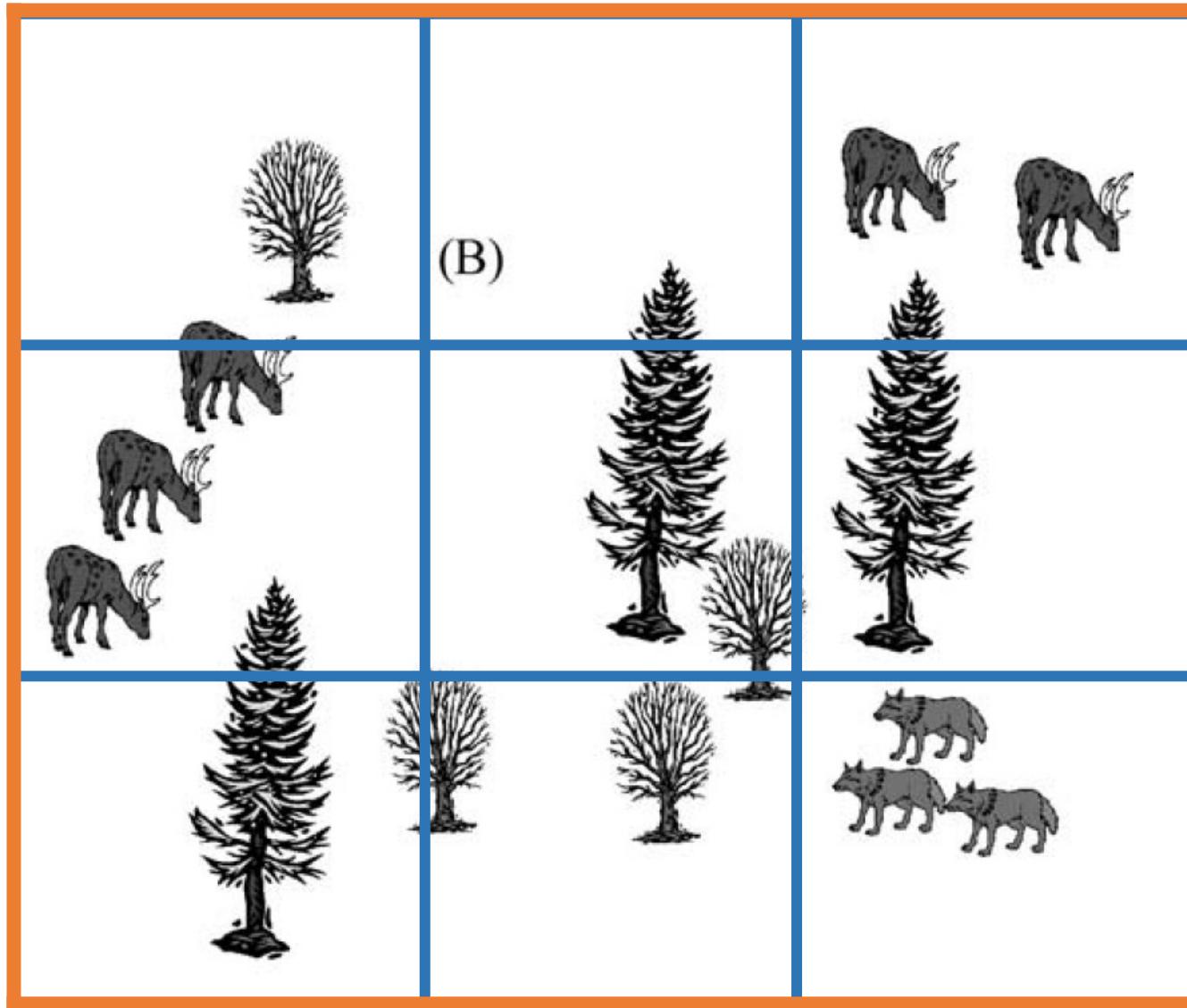
- ・人口は近年、増加している
 - > 全球スケール, 東京都スケール
- ・人口は近年、減少している
 - > 日本全国スケール、(地域スケール)



空間解像度(spatial grain/grain)が変わると？



空間解像度(spatial grain/resolution)



社会現象・自然現象とも規定する要因は、スケール(特に空間範囲)によって異なりうる

•気温と関係する要因の例

空間範囲	主要な要因
全球	気候帯
国	緯度(・標高)
都道府県/市町 村	標高、土地利用
字・町内会	地形

空間データを扱う際の落とし穴



- 「注目している現象」と「データ」のスケールの不一致
- 現象・現象を規定する要因のスケール依存性

空間データに関する基礎知識

- ✓ベクターデータとラスターデータ
- ✓座標参照(座標系と測地系)



Raster data and vector data

ラスターデータ:規則的に区切られたそれぞれの格子状区画（セル）が一つの値を持ち、空間情報をセルの集合体で表現



■**ベクターデータ:**点・線・面の3タイプ。位置や形状を頂点の座標で表現。一つの図形に複数の値を持たせることが可能



位置をどう表現?

OpenStreetMap

編集 履歴 エクスポート

GPSトレース 利用者の日記 コミュニティ 著作権 ヘルプ このサイトについて

東京スカイツリー 検索

検索結果 ✓ 東京都墨田区押上1丁目1-2

Results from OpenStreetMap Nominatim ✓ 35°71016288348506° N, 139°81072185771305° E
✓ X=12235283.7596165, Y=-6219635.287334526

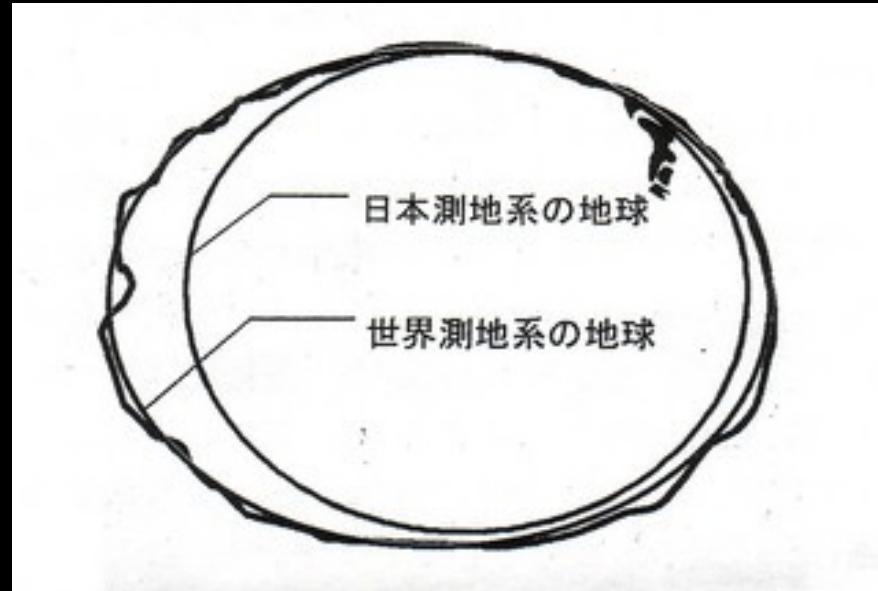
電波塔 東京スカイツリー, 2, 押上一丁目, 墨田区, 東京都, 131-0045, 日本

その他の結果

座標系

空間のある位置を、数字の組み合わせで表す際に、
その割り当て方を規定する規則

※座標の単位、原点の取り方、座標の順序等を規定



地球は完全な球状でなく、
赤道半径が極半径よりわずかに長い楕円球

測地系：

測地系は、この地球に似せた回転楕円体の
中心位置と赤道半径、扁平率などを定義

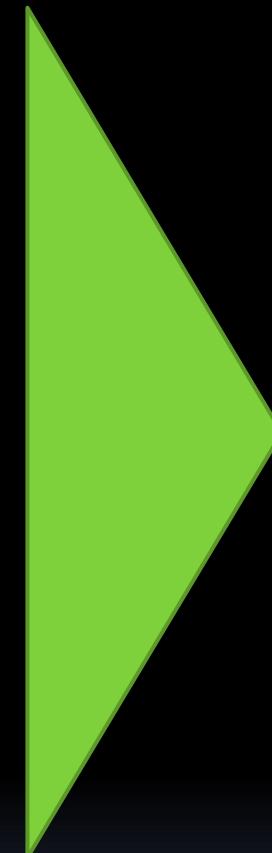
座標系 + 測地系 = 座標参照系(CRS: Coordinate Reference System)

日本でよく用いられる座標系

- ・平面直角座標系
- ・UTM（ユニバーサル横メルカトル）座標系
- ・経度緯度座標系

日本でよく用いられる測地系

- ・日本測地系2000 (Japanese Geodetic Datum 2000)
- ・日本測地系2011 (Japanese Geodetic Datum 2011)
- ・WGS84 (World Geodetic System 1984)



座標参照系
CRS

EPSGコード <https://epsg.io/>
(European Petroleum Survey Group)で識別

e.g.)
EPSG:6691 (JGD2011 UTM54N)
EPSG:4326 (WGS84 Lat/Lon)

最後に…

空間データサイエンスは魔法ではない

適切な

空間データ
サイエンス × •データ
 •該当分野の知識

大魔法使いへの道のりを楽しんで！

